

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 27

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5943 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(v i i) IMMEDIATE SOURCE:

(B) CLONE: Human Fibrinogen A-alpha chain

(i x) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(31..84, 1154..1279, 1739..1922, 3055..3200, 3786..5210)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

GTCTAGGAAC CAGCCCCACC CTTAGAAAAG ATG TTT TCC ATG AAG ATC GTC TGC      54
                                     Met Phe Ser Met Arg Ile Val Cys
                                     1               5

CTA GTT CTA AGT GTO GTO GGC ACA GCA TGG GTATGCCCT TTTCATTTTT      104
Leu Val Leu Ser Val Val Gly Thr Ala Trp
10               15

TCTTCTTGCT TTCTCTCTGG TGTATTATCC ACAAAGAGCC TGGAGGTCAO AGTCTACCTO      164

CTCTATGTCC TGACACACTC TTAOCTTTAT GACCCCAAGC CTGGGAGGAA ATTTCTCTGG      224

```

-continued

TGGGCTTGAC	ACCTCAAAGAA	TACAGGGTAA	TATGACACCA	AGAGGAAGAT	CTTAGATOGA	284
TGAGAGTGT	CAACTACAAG	GGAAACTTTA	GCATCTGTCA	TTCAGTCTTA	CCACATTTTG	344
TTTTGTTTTG	TTTTAAAAAG	GGCAAAGAATT	ATTTGCCATC	CTTGTACCTA	TAAAAGCCTTG	404
GTGCATTATA	ATGCTAGTTA	ATGGAATAAA	ACATTTTATG	GTAAGATTTG	TTTTCTTTAG	464
TTATTAATTT	CTTGCTACTT	GTCCATAATA	AGCAGAACTT	TTAGTGTTAAG	TACAGTTTTG	524
CTGAAAAGTT	ATTGTTGTGT	TTGTCAAGAC	AGAAGAAAAA	GCAAAACGAAT	TATCTTTTGA	584
AATATCTTTG	CAGTATCAGA	AGAGATTAGT	TAGTAAAGGA	ATACGCITTT	CCGCAGTAAAT	644
GGTATTCTTT	TAAATTATGA	ATCCATCTCT	AAAAGTTTACA	TAGAAAACCTG	AAGGAGAGAG	704
GAACATTCAAG	TTAAGATAAT	CTAAGTTTTT	CTACTGAAAG	AGCAATTACA	GGAGAAAAAG	764
CTCTACAATA	GTTTTCAACT	TTCTGTCTGC	AGTCATTAGT	AAAAATGAAA	AAGTAAAAATT	824
TAACTGATTT	TATAGATTCA	AATAATTTTC	CTTTTAGGAT	GGATTCTTTA	AAACTCCTAA	884
TATTTATCAA	ATGCTTATTT	AAGTGTACACA	CACAATTAAAG	AAATTTGTAC	ACCTTGTCTC	944
CITTAATTCT	CATAACAACT	CCATAAAATG	GGTCCTAGGA	TTTCCATTTG	AAGATAAGAA	1004
ACCTGAAAGCT	TGCCGAAAGCC	CTGTGTCTGC	TCTCCTTAAT	CTCTGTGAGA	GTGCCATCTC	1064
TTCTTGGGGA	CTTGTAAAGCA	TGCCACTGTC	TCCTCTTCTG	GCTAACATTG	CTGTTGCTCT	1124
CTTTTGTGTA	TGTGAATGAA	TCTTTAAAG	ACT GCA GAT	AGT GGT GAA	GGT GAC	1177
			Thr Ala Asp Ser Gly Glu Gly Asp			
			20		25	
TTT CTA GCT	GAA GGA GGA GGC	GTG COT GGC	CCA AGG GTT	GTG GAA AGA		1225
Phe Leu Ala	Glu Gly Gly Gly Val	Arg Gly Pro	Arg Val Val	Glu Arg		
	30	35		40		
CAT CAA TCT	GCC TGC AAA GAT	TCA GAC TGG	CCC TTC TGC	TCT GAT GAA		1273
His Glu Ser	Ala Cys Lys Asp	Ser Asp Trp	Pro Phe Cys	Ser Asp Glu		
	45	50		55		
GAC TGG GTAAACAGTC	AACGAGGAGAA	GCAGGAGATT	CCTTCCCTCT	GATGCTAGAG		1329
Asp Trp						
	60					
GGGCTCACAG	GCTGACCTGA	TTGGTCCCAAG	AAACTTTTTT	AAATAGAAAA	TAATTGAATA	1389
GTTACCTACA	TAACAAATAA	AGAAAAAGAA	CCTACTCCCA	AGAGCACTGT	TTATTTACCT	1449
CCCCAACTCT	GGATCATTAG	TGGGTGAACA	GACAGGATTT	CAGTTGCATG	CTCAGGCAAA	1509
ACCAGGCTCC	TGAGTATTGT	GGCCTCAATT	TCCTGGCACC	TATTTATGGC	TAAATGGACC	1569
CTCATTCCAAG	AGTTTCTCTG	CGACCTCTAA	CTAGTCCTCT	TACCTACTTT	TAAACCAACT	1629
TATCTGGAAG	AGAAAAAGGTA	GGAAAGAAATG	GGGGCTGCAT	GGAAACATGC	AAAAATTATTC	1689
TGAATCTGAG	AGATAGATCC	TTACTGTAAT	TTTCTCCCTT	CACTTTCAG	AAC TAC	1744
AAA TGC CCT	TCT GGC TGC	AGG ATG AAA	GGG TTG ATT	GAT GAA	GTC AAT	1792
Lys Cys Pro	Ser Gly Cys Arg	Met Lys Gly Leu	Ile Asp Glu	Val Asn		
	65	70		75		
CAA GAT TTT	ACA AAC AGA	ATA AAT AAG	CTC AAA AAT	TCA CTA	TTT GAA	1840
Glu Asp Phe	Thr Asn Arg	Ile Asn Lys	Leu Lys Asn	Ser Leu Phe	Glu	
	80	85		90		
TAT CAG AAG	AAC AAT AAG	GAT TCT CAT	TCG TTG ACC	ACT AAT	ATA ATG	1888
Tyr Glu Lys	Asn Asn Lys	Asp Ser His	Ser Leu Thr	Thr Thr	Asn Ile Met	
	95	100		105	110	
GAA ATT TTG	AGA GGC GAT	TTT TCC TCA	GCC AAT A	GTAAATATTA		1932
Glu Ile Leu	Arg Gly Asp Phe	Ser Ser Ala	Asn			
	115	120				
CATATTTACT	TCTTTGACTT	TATAACAGAA	ACAACAAAAA	TCCTAAATAA	ATATGATATC	1992

-continued

CGCTTATATC	TATGACAATT	TCATCCCAAA	GTACTTAGTG	TAGAAAACACA	TACCTTCATA	2052
ATATCCCTGA	AAATTTTAAO	AOGGAGCTTT	TGTTTTCGTT	ATTTTTTCAA	AGTAAAAAGAT	2112
GTTAACTGAG	ATTGTTTAAO	OTCACAATAAT	AAGTCAGAAT	TTTGGATTAA	AACAAGGAATT	2172
TAAATGTGTT	CTTTTCAACA	GTATATACTG	AAAAGTAGGAT	GGGTCAGACT	CTTTGAGTTG	2232
ATATTTTTGT	TTCTGCTTTG	TAAAGGTGAA	AACTGAGAGG	TCAAAGAACT	TGTTCAAAGA	2292
CACAGAGCTG	GGAATTCAAC	TCCCAGACTC	CAGTGAAGCTG	ATTAGGTAGA	TTTTTAAATT	2352
TAAAAATATG	GGTCAAGCTA	COTCAATTCTC	ACAGTCTACT	CATTAGGGTT	AGGAAACATT	2412
GCATTCACCTC	TGGGCATOGA	CAOGCAAGTCT	AOGGAGTCCCT	CAGTTTCTCA	AGTTTTGCTT	2472
TGCTTTTTTA	CACCTTCACA	AACACTTGAC	ATTTAAAAATC	AGTGATGCCA	ACACTAGCTG	2532
GCAAGTGAAT	GATCCTGTTG	ACCCAAAAACA	GCTTAGGAAC	CATTTCAAAT	CTATAGAGTT	2592
AAAAAGAAAA	GCTCATCAAT	AAGAAAAATCC	AATATGTTCA	AGTCCCTTGA	TAAAGGATGT	2652
TATAAAATAA	TGAAATOGA	ATCAAAACCA	CTATTTTAAAC	TCCAAATTAC	ACCTTTAAAA	2712
TTCCAAAGAA	AGTTCTTCTT	CTATATTTCT	TTGGGATTAC	TAATTGCTAT	TAGGACATCT	2772
TAACTGGCAT	TCATGGAAAG	CTGCAAGGCA	TAACATTATC	CAAAAGTCAA	ATGCCCCATA	2832
GGTTTTGAAC	TCACAGATTA	AACTGTAAAC	AAAAATAAAAT	TAGGCATATT	TACAAAGCTAG	2892
TTTTCTTCTT	TCTTTTTCTT	CTTTCTTCTT	TTCTTTCTTT	CTTTCTTCTT	TTCTTTCTTT	2952
CTTTCTTCTT	TTCTCCTTCC	TTCTTTCTT	CCTTTCTTTT	TTGCTGGCAA	TTACAGACAA	3012
ATCACTCAGC	AGCTACTTCA	ATAACCATAT	TTTCGATTTT	AG	AC COT GAT AAT	3065
Asn Arg Asp Asn 125						
ACC TAC AAC CGA GTO TCA GAG GAT CTO AGA AOC	AGA ATT GAA GTC CTO	3113				
Thr Tyr Asn Arg Val Ser Glu Asp Leu Arg Ser Arg Ile Glu Val Leu						
130	135	140				
AAG CGC AAA GTC ATA GAA AAA GTA CAG CAT ATC	CAG CTT CTO CAG AAA	3161				
Lys Arg Lys Val Ile Glu Lys Val Glu His Ile Glu Leu Leu Glu Lys						
145	150	155				
AAT GTT AOA OCT CAG TTO GTT GAT ATG AAA COA	CTG GAG GTAAATATGT	3210				
Asn Val Arg Ala Glu Leu Val Asp Met Lys Arg Leu Glu						
160	165	170				
GGCTGTGCTC	CCGAGTGCTC	TTOTTTTTGA	GTAGAGGGGAA	AAAGAAAGGCG	ATAATTATGC	3270
ACTGAGTGTC	TACTATATGC	AGAGAAAAAT	OTTATATCCA	TCATCTACCT	AAAAAGTAAGT	3330
ATTATTTTTCC	TCACTCCACA	GTTOAAAGAAA	AAAAAATICA	GAGATATTAA	GTAAATTTTC	3390
CAACGTACAT	AGATAGTAAT	TCAAAOCAAT	GTTCAGTCCC	TGTCTATTCC	AAOCCATTAC	3450
ATCACCACAC	CTCTGAGCCC	TCAOGCTGAG	TTACCCAAAG	ATCATTTAAT	TAGCGTTTCC	3510
TTTGAAGAGG	AATAGCACCT	TACTCTTGAT	CCATTCTGAG	GCTAAGATGA	ATTAAACAGC	3570
ATCCATTGCT	TATCCTGGCT	AGCCCTGCAA	TACCCAACAT	CTCTTCCACT	GAGGGTGCTC	3630
GATAGGCAOA	AAACAGAGAA	TATTAAAGTG	TAGGTCTCCG	AGTCAAAAAA	AATGAAACCA	3690
GTTTCCAGAA	GGAAAAATTAA	CTACCAAGAA	CTCAATAGAC	GTAGTTTATG	TATTTGTATC	3750
TACATTTTCT	CTTTATTTTT	CTCCCTCTC	TCTAG GTG GAC	ATT GAT ATT AAG		3803
Val Asp Ile Asp Ile Lys 175						
ATC COA TCT TGT COA GGG TCA TGC AGT AOG OCT	TTA OCT COT GAA GTA	3851				
Ile Arg Ser Cys Arg Gly Ser Cys Ser Arg Ala Leu Ala Arg Glu Val						
180	185	190				
GAT CTG AAG GAC TAT GAA GAT CAG CAG AAG CAA	CTT GAA CAG GTC ATT	3899				

-continued

Asp	Leu	Lys	Asp	Tyr	Glu	Asp	Glu	Glu	Lys	Glu	Leu	Glu	Glu	Val	Ile	
		195					200					205				
GCC	AAA	GAC	TTA	CTT	CCC	TCT	AGA	GAT	AGG	CAA	CAC	TTA	CCA	CTG	ATA	3947
Ala	Lys	Asp	Leu	Leu	Pro	Ser	Arg	Asp	Arg	Gln	His	Leu	Pro	Leu	Ile	
	210					215					220					
AAA	ATG	AAA	CCA	GTT	CCA	GAC	TGG	GTT	CCC	GGA	AAT	TTT	AAG	AGC	CAG	3995
Lys	Met	Lys	Pro	Val	Pro	Asp	Leu	Val	Pro	Gly	Asn	Phe	Lys	Ser	Gln	
	225				230					235					240	
CTT	CAG	AAG	GTA	CCC	CCA	GAG	TGG	AAG	GCA	TTA	ACA	GAC	ATG	CCG	CAG	4043
Leu	Gln	Lys	Val	Pro	Pro	Gln	Trp	Lys	Ala	Leu	Thr	Asp	Met	Pro	Gln	
			245						250					255		
ATG	AGA	ATG	GAG	TTA	GAG	AGA	CCT	GGT	GGA	AAT	GAG	ATT	ACT	CGA	GGA	4091
Met	Arg	Met	Glu	Leu	Glu	Arg	Pro	Gly	Gly	Asn	Glu	Ile	Thr	Arg	Gly	
			260					265					270			
GGC	TCC	ACC	TCT	TAT	GGA	ACC	GGA	TCA	GAG	ACG	GAA	AGC	CCC	AGG	AAC	4139
Gly	Ser	Thr	Ser	Tyr	Gly	Thr	Gly	Ser	Gln	Thr	Glu	Ser	Pro	Arg	Asn	
		275					280					285				
CCT	AGC	AGT	GCT	GGA	AGC	TGG	AAC	TCT	GGG	AGC	TCT	GGA	CCT	GGA	AGT	4187
Pro	Ser	Ser	Ala	Gly	Ser	Trp	Asn	Ser	Gly	Ser	Ser	Gly	Pro	Gly	Ser	
	290					295						300				
ACT	GGA	AAC	CGA	AAC	CCT	GGG	AGC	TCT	GGG	ACT	GGA	GGG	ACT	GCA	ACC	4235
Thr	Gly	Asn	Arg	Asn	Pro	Gly	Ser	Ser	Gly	Thr	Gly	Gly	Thr	Ala	Thr	
	305				310					315					320	
TGG	AAA	CCT	GGG	AGC	TCT	GGA	CCT	GGA	AGT	GCT	GGA	AGC	TGG	AAC	TCT	4283
Trp	Lys	Pro	Gly	Ser	Ser	Gly	Pro	Gly	Ser	Ala	Gly	Ser	Trp	Asn	Ser	
			325						330					335		
GGG	AGC	TCT	GGA	ACT	GGA	AGT	ACT	GGA	AAC	CAA	AAC	CCT	GGG	AGC	CCT	4331
Gly	Ser	Ser	Gly	Thr	Gly	Ser	Thr	Gly	Asn	Gln	Asn	Pro	Gly	Ser	Pro	
			340					345					350			
AGA	CCT	GGT	AGT	ACC	GGA	ACC	TGG	AAT	CCT	GGC	AGC	TCT	GAA	CGC	GGA	4379
Arg	Pro	Gly	Ser	Thr	Gly	Thr	Trp	Asn	Pro	Gly	Ser	Ser	Gln	Arg	Gly	
		355					360					365				
AGT	GCT	GGG	CAC	TGG	ACC	TCT	GAG	AGC	TCT	GTA	TCT	GGT	AGT	ACT	GGA	4427
Ser	Ala	Gly	His	Trp	Thr	Ser	Glu	Ser	Ser	Val	Ser	Gly	Ser	Thr	Gly	
	370					375						380				
CAA	TGG	CAC	TCT	GAA	TCT	GGA	AGT	TTT	AGG	CCA	GAT	AGC	CCA	GGC	TCT	4475
Gln	Trp	His	Ser	Gln	Ser	Gly	Ser	Phe	Arg	Pro	Asp	Ser	Pro	Gly	Ser	
	385				390					395					400	
GGG	AAC	GGC	AGG	CCT	AAC	AAC	CCA	GAC	TGG	GGC	ACA	TTT	GAA	GAG	GTO	4523
Gly	Asn	Ala	Arg	Pro	Asn	Asn	Pro	Asp	Trp	Gly	Thr	Phe	Gln	Glu	Val	
			405						410					415		
TCA	GGA	AAT	GTA	AGT	CCA	GGG	ACA	AGG	AGA	GAG	TAC	CAC	ACA	GAA	AAA	4571
Ser	Gly	Asn	Val	Ser	Pro	Gly	Thr	Arg	Arg	Gln	Tyr	His	Thr	Gln	Lys	
			420					425					430			
CTG	GTC	ACT	TCT	AAA	GGA	GAT	AAA	GAG	CTC	AGG	ACT	GGT	AAA	GAG	AAG	4619
Leu	Val	Thr	Ser	Lys	Gly	Asp	Lys	Gln	Leu	Arg	Thr	Gly	Lys	Gln	Lys	
		435					440					445				
GTC	ACC	TCT	GGT	AGC	ACA	ACC	ACC	ACG	CGT	CGT	TCA	TGC	TCT	AAA	ACC	4667
Val	Thr	Ser	Gly	Ser	Thr	Thr	Thr	Thr	Arg	Arg	Ser	Cys	Ser	Lys	Thr	
		450				455					460					
GTT	ACT	AAG	ACT	GTT	ATT	GGT	CCT	GAT	GGT	CAC	AAA	GAA	GTT	ACC	AAA	4715
Val	Thr	Lys	Thr	Val	Ile	Gly	Pro	Asp	Gly	His	Lys	Gln	Val	Thr	Lys	
	465				470					475					480	
GAA	GTO	GTO	ACC	TCC	GAA	GAT	GGT	TCT	GAC	TGT	CCC	GAG	GCA	ATG	GAT	4763
Glu	Val	Val	Thr	Ser	Gln	Asp	Gly	Ser	Asp	Cys	Pro	Gln	Ala	Met	Asp	
			485						490					495		
TTA	GGC	ACA	TTO	TCT	GGC	ATA	GGT	ACT	CTG	GAT	GGG	TTC	CGC	CAT	AGG	4811
Leu	Gly	Thr	Leu	Ser	Gly	Ile	Gly	Thr	Leu	Asp	Gly	Phe	Arg	His	Arg	
			500					505					510			
CAC	CCT	GAT	GAA	GCT	GCC	TTC	TTC	GAC	ACT	GCC	TCA	ACT	GGA	AAA	ACA	4859

-continued

His	Pro	Asp	Glu	Ala	Ala	Phe	Phe	Asp	Thr	Ala	Ser	Thr	Gly	Lys	Thr	
		515					520					525				
TTC	CCA	GGT	TTC	TTC	TCA	CCT	ATG	TTA	GGA	GAG	TTT	GTC	AGT	GAG	ACT	4907
Phe	Pro	Gly	Phe	Phe	Ser	Pro	Met	Leu	Gly	Glu	Phe	Val	Ser	Glu	Thr	
	530					535					540					
GAG	TCT	AGG	GGC	TCA	GAA	TCT	GGC	ATC	TTC	ACA	AAT	ACA	AAG	GAA	TCC	4955
Glu	Ser	Arg	Gly	Ser		Ser	Gly	Ile	Phe		Asa	Thr	Lys	Glu	Ser	
	545				550				555					560		
AGT	TCT	CAT	CAC	CCT	GGG	ATA	GCT	GAA	TTC	CCT	TCC	CGT	GGT	AAA	TCT	5003
Ser	Ser	His	His	Pro	Gly	Ile	Ala	Glu	Phe	Pro	Ser	Arg	Gly	Lys	Ser	
				565				570						575		
TCA	AGT	TAC	AGC	AAA	CAA	TTT	ACT	AGT	AGC	ACG	AGT	TAC	AAC	AGA	GGA	5051
Ser	Ser	Tyr	Ser	Lys	Gln	Phe	Thr	Ser	Ser	Thr	Ser	Tyr	Asa	Arg	Gly	
			580					585					590			
GAC	TCC	ACA	TTT	GAA	AOC	AAO	AGC	TAT	AAA	ATG	GCA	GAT	GAG	GCC	GGA	5099
Asp	Ser	Thr	Phe	Glu	Ser	Lys	Ser	Tyr	Lys	Met	Ala	Asp	Glu	Ala	Gly	
	595						600					605				
AGT	GAA	GCC	GAT	CAT	GAA	GGA	ACA	CAT	AGC	ACC	AAO	AGA	GGC	CAT	GCT	5147
Ser	Glu	Ala	Asp	His	Glu	Gly	Thr	His	Ser	Thr	Lys	Arg	Gly	His	Ala	
	610					615					620					
AAA	TCT	CGC	CCT	GTC	AGA	GGT	ATC	CAC	ACT	TCT	CCT	TTG	GGG	AAO	CCT	5195
Lys	Ser	Arg	Pro	Val	Arg	Gly	Ile	His	Thr	Ser	Pro	Leu	Gly	Lys	Pro	
	625				630					635				640		
TCC	CTG	TCC	CCC	TAGACTAAGT	TAAATATTTC	TOCACAGTGT	TCCCATGCCC									5247
Ser	Leu	Ser	Pro													
				645												
CCTTGCAATTT	CCTTCTTAAC	TCTCTGTTAC	ACGTCAATGA	AACTACACTT	TTTTGGTCTG											5307
TTTTTGTGCT	AGACTGTAAO	TTCCTTGGGG	GCAOGGOCCTT	TGTCTGTCTC	ATCTCTGTAT											5367
TCCCAAATOC	CTAACAGTAC	AGAGCCATGA	CTCAATAAAT	ACATGTTAAA	TGGATGAATO											5427
AATTCCTCTG	AAACTCTATT	TGAGCTTATT	TAOTCAAATT	CTTTCACTAT	TCAAAGTGTG											5487
TGCTATTAGA	ATTGTCACCC	AACTGATTAA	TCACATTTTT	AGTATGTGTC	TCAOTTGACA											5547
TTTAGGTCAO	GCTAAATACA	AGTTGTGTTA	GTATTAAOTG	AGCTTAGCTA	CCTGTACTOG											5607
TTACTTGCTA	TTAGTTTGTG	CAAGTAAAAT	TCCAAATACA	TTTGAAGAAA	ATCCCCTTTG											5667
CAATTTGTAG	GTATAAATAA	CCGCTTATTT	GCATAAGTTC	TATCCCCTG	TAAOTOCATC											5727
CTTTCCTTAT	GGAGGGAAAG	AAAGGAGGAA	GAAAGAAAAG	AAAGGAAAAG	AACAGTATTT											5787
GCCTTATTTA	ATCTGAGCCG	TGCCTATCTT	TGTAAAGTTA	AATGAGAATA	ACTTCTTCCA											5847
ACCAOCTTAA	TTTTTTTTTT	AGACTGTGAT	GATGTCCTCC	AAACACATCC	TTCAOGTACC											5907
CAAAGTGGCA	TTTTCATAT	CAAGCTATCC	GGATCC													5943

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 644 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Phe Ser Met Arg Ile Val Cys Leu Val Leu Ser Val Val Gly Thr
1 5 10 15

Ala Trp Thr Ala Asp Ser Gly Glu Gly Asp Phe Leu Ala Glu Gly Gly
20 25 30

Gly Val Arg Gly Pro Arg Val Val Glu Arg His Gln Ser Ala Cys Lys
35 40 45

-continued

Asp	Ser	Asp	Trp	Pro	Phe	Cys	Ser	Asp	Glu	Asp	Trp	Asa	Tyr	Lys	Cys	
50						55					60					
Pro	Ser	Gly	Cys	Arg	Met	Lys	Gly	Leu	Ile	Asp	Glu	Val	Asa	Gln	Asp	
65					70					75					80	
Phe	Thr	Asa	Arg	Ile	Asa	Lys	Leu	Lys	Asa	Ser	Leu	Phe	Glu	Tyr	Gln	
				85					90					95		
Lys	Asa	Asa	Lys	Asp	Ser	His	Ser	Leu	Thr	Thr	Asa	Ile	Met	Glu	Ile	
			100					105					110			
Leu	Arg	Gly	Asp	Phe	Ser	Ser	Ala	Asa	Asa	Arg	Asp	Asa	Thr	Tyr	Asa	
		115					120					125				
Arg	Val	Ser	Glu	Asp	Leu	Arg	Ser	Arg	Ile	Glu	Val	Leu	Lys	Arg	Lys	
	130					135					140					
Val	Ile	Glu	Lys	Val	Gln	His	Ile	Gln	Leu	Leu	Gln	Lys	Asa	Val	Arg	
145					150					155					160	
Ala	Gln	Leu	Val	Asp	Met	Lys	Arg	Leu	Glu	Val	Asp	Ile	Asp	Ile	Lys	
				165					170					175		
Ile	Arg	Ser	Cys	Arg	Gly	Ser	Cys	Ser	Arg	Ala	Leu	Ala	Arg	Glu	Val	
			180					185					190			
Asp	Leu	Lys	Asp	Tyr	Glu	Asp	Gln	Gln	Lys	Gln	Leu	Gln	Gln	Val	Ile	
		195					200					205				
Ala	Lys	Asp	Leu	Leu	Pro	Ser	Arg	Asp	Arg	Gln	His	Leu	Pro	Leu	Ile	
	210					215					220					
Lys	Met	Lys	Pro	Val	Pro	Asp	Leu	Val	Pro	Gly	Asa	Phe	Lys	Ser	Gln	
225					230					235					240	
Leu	Gln	Lys	Val	Pro	Pro	Glu	Trp	Lys	Ala	Leu	Thr	Asp	Met	Pro	Gln	
				245					250					255		
Met	Arg	Met	Glu	Leu	Glu	Arg	Pro	Gly	Gly	Asa	Glu	Ile	Thr	Arg	Gly	
			260					265					270			
Gly	Ser	Thr	Ser	Tyr	Gly	Thr	Gly	Ser	Glu	Thr	Glu	Ser	Pro	Arg	Asa	
		275					280					285				
Pro	Ser	Ser	Ala	Gly	Ser	Trp	Asa	Ser	Gly	Ser	Ser	Gly	Pro	Gly	Ser	
	290					295					300					
Thr	Gly	Asa	Arg	Asa	Pro	Gly	Ser	Ser	Gly	Thr	Gly	Gly	Thr	Ala	Thr	
305					310					315					320	
Trp	Lys	Pro	Gly	Ser	Ser	Gly	Pro	Gly	Ser	Ala	Gly	Ser	Trp	Asa	Ser	
				325					330					335		
Gly	Ser	Ser	Gly	Thr	Gly	Ser	Thr	Gly	Asa	Gln	Asa	Pro	Gly	Ser	Pro	
			340					345					350			
Arg	Pro	Gly	Ser	Thr	Gly	Thr	Trp	Asa	Pro	Gly	Ser	Ser	Glu	Arg	Gly	
		355					360					365				
Ser	Ala	Gly	His	Trp	Thr	Ser	Glu	Ser	Ser	Val	Ser	Gly	Ser	Thr	Gly	
	370					375					380					
Gln	Trp	His	Ser	Glu	Ser	Gly	Ser	Phe	Arg	Pro	Asp	Ser	Pro	Gly	Ser	
385					390					395					400	
Gly	Asa	Ala	Arg	Pro	Asa	Asa	Pro	Asp	Trp	Gly	Thr	Phe	Glu	Glu	Val	
				405					410					415		
Ser	Gly	Asa	Val	Ser	Pro	Gly	Thr	Arg	Arg	Glu	Tyr	His	Thr	Glu	Lys	
			420					425					430			
Leu	Val	Thr	Ser	Lys	Gly	Asp	Lys	Glu	Leu	Arg	Thr	Gly	Lys	Glu	Lys	
		435					440					445				
Val	Thr	Ser	Gly	Ser	Thr	Thr	Thr	Thr	Arg	Arg	Ser	Cys	Ser	Lys	Thr	
	450					455					460					
Val	Thr	Lys	Thr	Val	Ile	Gly	Pro	Asp	Gly	His	Lys	Glu	Val	Thr	Lys	
465					470					475					480	

-continued

Glu Val Val Thr Ser Glu Asp Gly Ser Asp Cys Pro Glu Ala Met Asp
 485 490 495
 Leu Gly Thr Leu Ser Gly Ile Gly Thr Leu Asp Gly Phe Arg His Arg
 500 505 510
 His Pro Asp Glu Ala Ala Phe Phe Asp Thr Ala Ser Thr Gly Lys Thr
 515 520 525
 Phe Pro Gly Phe Phe Ser Pro Met Leu Gly Glu Phe Val Ser Glu Thr
 530 535 540
 Glu Ser Arg Gly Ser Glu Ser Gly Ile Phe Thr Asn Thr Lys Glu Ser
 545 550 555 560
 Ser Ser His His Pro Gly Ile Ala Glu Phe Pro Ser Arg Gly Lys Ser
 565 570 575
 Ser Ser Tyr Ser Lys Glu Phe Thr Ser Ser Thr Ser Tyr Asn Arg Gly
 580 585 590
 Asp Ser Thr Phe Glu Ser Lys Ser Tyr Lys Met Ala Asp Glu Ala Gly
 595 600 605
 Ser Glu Ala Asp His Glu Gly Thr His Ser Ser Lys Arg Gly His Ala
 610 615 620
 Lys Ser Arg Pro Val Arg Gly Ile His Thr Ser Pro Leu Gly Lys Pro
 625 630 635 640
 Ser Leu Ser Pro

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8878 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

- (B) CLONE: human Rhinogen B-beta chain

(ix) FEATURE:

- (A) NAME/KEY: misc_RNA
- (B) LOCATION: 1..469

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 470..583

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 584..3257

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 3258..3469

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 3450..3938

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 3939..4127

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 4123..5042

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 5043..5270

(ix) FEATURE:

-continued

(A) NAME/KEY: intron
(B) LOCATION: 5271..5830

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 5831..5944

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 5945..6632

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 6633..6758

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 6759..6966

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 6967..7252

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 7253..7870

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 7871..8102

(ix) FEATURE:
(A) NAME/KEY: 3'UTR
(B) LOCATION: 8103..8537

(ix) FEATURE:
(A) NAME/KEY: misc_RNA
(B) LOCATION: 8538..8878

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: join(470..583, 5258..3449, 3939..4122,
5043..5270, 5831..5944, 6633..6758, 6967..7252,
7871..8102)

(xi) SEQUENCE DESCRIPTION: SBQ ID NO3:

```

GAATTCATGC CCCTTTTGAA ATAGACTTAT GTCATTGTCA GAAAACATAA GCATTTATGG      60
TATATCATT AATGATCACG ATTTTAGTGG TTGCTTTGTG AGTAGGTCAA ATTTACTAAG      120
CTTAGATTGG TTTTCTCACA TATTCCTTCG GAGCTTGTGT AGTTTCCACA TTAATTTACC      180
AGAAACAAGA TACACACTCT CTTTGAAGAG TCCCTAACT TCCCATCATT TTGTCCAATT      240
AAATGAATTG AAGAAATTTA ATGTTTCTAA ACTAGACCAA CAAAGAATAA TAGTTGTATG      300
ACAAGTAAAT AAGCTTTGCT GGGAAAGATGT TGCTTAAATG ATAAAATGGT TCAGCCAACA      360
AGTGAACCAA AAATTAAATA TTAAGTAAAG AAAAGTAACC ATTTCTGAAG TCATTCTAG      420
CAGAGGACTC AGATATATAT AGGATTGAAG ATCTCTCAGT TAACTCTAC ATG AAA          475

                                     Met Lys
                                     1

AGG ATG GTT TCT TGG AOC TTC CAC AAA CTT AAA ACC ATG AAA CAT CTA      523
Arg Met Val Ser Trp Ser Phe His Lys Leu Lys Thr Met Lys His Leu
                                     5          10          15

TTA TTG CTA CTA TTG TGT GTT TTT CTA GTT AAG TCC CAA GGT GTC AAC      571
Leu Leu Leu Leu Leu Cys Val Phe Leu Val Lys Ser Glu Gly Val Asn
                                     20          25          30

GAC AAT GAG GAG GTGAATTTTT TAAAGCATT TATATTATT AGTAGTATTA      623
Asp Asn Glu Glu
                                     35

TTAATATAAG ATGTAACATA ATCATATTAT GTGCTTATTT TAATGAAATT AGCATTGCTT      683

```


-continued

ATAOTTATGA	AATOGAATTO	TAAACCTCTG	ACTTATTGTA	TTTAAAGAAT	GTTTCATAOT	743
ATTTCTTATA	TAAAAACAAA	OTAATTTCTT	GTTTTCTAGT	TTATCACCTT	TGTTTTCTTA	803
AGATGAOGAT	GGCTTAGCTA	ATGTAAOATG	TGTTTTTCTC	ACTTGCTATT	CTOAGTACTG	863
TGATTTTCAT	TTACTTCTAG	CAATACAGGA	TTACAATTAA	GAOGACAAGA	TCTGAAAAATC	923
TCACAAACTA	TAAAAATAATA	AAAGAOCAGA	ATTTTAAAGAT	AAAAAGAAACT	GGTGGTAGGT	983
AGATTGTTCT	TTGGTGAAAG	AAOGTAATAT	ATATTGTTAC	TGAGATTACT	ATTTATAAAAA	1043
ATTATAACTA	AGCCTAAAAG	CAAAATACAT	CAAGTGTAAAT	GATAGAAAAAT	GAAATATTGC	1103
TTTTTTTCAOA	TGAAAAAGTTC	AAATTAOAGT	TAOTGTGTAT	TGTTATTATT	AATAOTTATG	1163
AAACACOGTT	CAGTCTAATT	TATTTATTTG	TAGAACAGTT	TGTCCTCAAC	TATTTATTTT	1223
GCTGACTTAT	TGCTGTAAAT	TTGCAOTTAC	TAAAAATACA	GAAATGCATT	TAOGACAATG	1283
GATATTTAAG	AAATTTAAAT	TTTATCATCA	AACGTATCAT	GGCCAAATTT	CTTACATATA	1343
GCATAGTATC	ATTAAACTAG	AAATAAGAAAT	ACACAATAAT	ATTTAAATGA	AGTGATTTCAT	1403
TTGCGATCAT	TATTGAGTTT	CAAOGGAACT	TGAGTGTGTT	ACTTATCAGA	CTCTACATGT	1463
AAGAACATAT	AGTTAATCTG	GTGTGTGTGT	TAAAAACATA	TGGTTAATCT	GGTTAAOTCT	1523
GGTTAATCAT	ATTAGGTAAO	AAAAATGTAA	AGAAATGTGT	AGACGAAATT	TTTGTAAGGT	1583
ACTCTGCAAA	GCACCTTCAC	ATTTCTGCTT	ATCAACTAAA	CCTCACAGAG	ATAOTTTAAT	1643
AGTTTAGOCT	TTAAAAATGA	TTTTGATTAT	TCAACAAOTG	GCCTTCATAA	TTTCTTTAAG	1703
TGTTTTTCTT	TAGTATATATA	CTTTCTTTAA	ATATTTTTTA	AAATTTCCCT	TTCTCTAGTA	1763
AAGCCAGACC	ATCCATGCTA	CCTCTCTAGT	GGCACTCTGA	AATAAAAAAG	AAATAOTTTT	1823
CTCTGTTATA	ATTGTATTTG	TAATAAGCAG	ATGAATCACA	TTTCTTAAAA	TTTGTTTTAG	1883
AGAGGGTAAO	CTCTGACTAG	GACCATGACT	TCAATGTGAA	ATATGTATAT	ATCCTCCGAA	1943
TCTTTACATA	TTAAGAATGT	ATATAGTCAA	CTGGTTAAAC	AGGAAAAATCT	GGAAACAOCCT	2003
GGCTGGGTTT	TAATCTTAGC	ACCATCCTAC	TAAATGTTAA	ATAATATTAT	AATCTAATGA	2063
ATAAATGACA	ATGCAATTCC	AAATAOAGTT	CATCTOATGA	CTTCTAGACT	CACAAAAATG	2123
CAAGAGAGCT	CAGTTGTTGC	TCAOTTGTTT	CAAAATCATGT	CGTTTGTTAA	TTTGTAATTA	2183
AGCTCCAAAO	GATGTATAGC	TACTGACAAA	AAAAAAAAATG	AGAAATGTAGT	TAATCCAAAT	2243
CAAAACTTTT	CTATTGCAAT	GCOTATTTTC	TGCTTCATTA	TCCTTTAATA	TAATATTTTA	2303
AGTTAGCAAG	TAATTTTAAAT	TACAATGCAC	AAOCCTTGAG	AATTATTTTA	AATATAAGAA	2363
AATCATAATG	TTTGATAAAO	AAATCATGTA	AGAAATTTCA	AGATAATGGT	TTAACAAATA	2423
ATTTTGTTGA	TAGAAAGATA	GACTAAAAAT	GAAATTCGAA	GTGGAGAGGA	CACTTAAACT	2483
GTAGTACTTG	TTATGTGTGA	TTCCAOTAAA	AATAOTAAAT	AOCATTATT	ATTOCCAAAT	2543
ACTGTTCTGA	GGGTACCATA	TGCAATAAAT	TATTTAATCC	TTACAATAAT	CTTGTAAGGC	2603
AGATTCAAAAC	TATCATTACA	CTTATTTTAC	AGATGAGAAA	ACTGGGGCAC	AGATAAAOCA	2663
ACTTGCCCAA	GGTCTCATAG	CTGTAAATCA	ACCCTACGGT	CAAGACCTAC	AAOTAGCCGA	2723
GCTCCAGAGT	ACATTATGAG	GGTCAAAAGAT	TGTCTTATTA	CAAAATAAAT	CCAAOTAGAA	2783
TCAACCTTTA	ATAAGTCTTT	AATGTCTCTT	AAATATGTTT	ATATAOGAAT	CTAATCACCA	2843
ATTACAAAAA	ATGAAAGTAG	GGAAATGATT	AACAATAATC	ATAGGAATCT	AACAATCCAA	2903
GTGGCTTGAO	AATATTTCATT	CTTCTTGACA	GTATAOATTC	TTTACAATTT	CGTAAOTTCC	2963
AATGTATGTT	TTAGGAATAT	GAAGTCATTA	CTATTCATAA	TCTOATACAG	CTTTATCCTA	3023
AGGCCTCTCT	TTAAAAACTA	CACGTGCATCA	TAOCTTTTTT	GTGCAOTTGG	TCTTTCTACT	3083

-continued

GTTACTGAAC	AGTAAOCAAC	CTACAGATTC	ACTATCACCA	ACCAOCCAAT	TGATGGATCT	3143
TAAGCAAAAT	ATCAAGCTTG	TGATAACCTA	AATTATAAAA	TGAGGGTGTT	GGAAATAGTTA	3203
CATTCCAAAT	CTTCTATAAC	ACTCTGTATT	ATATTTCTGC	CTCATTCCCT	GTAG GGT	3260
					Gly	
TTC TTC AGT GCC CGT GGT CAT CGA CCC CTT GAC AAG AAG AGA GAA GAG	3308					
Phe Phe Ser Ala Arg Gly His Arg Pro Leu Asp Lys Lys Arg Glu Glu						
40 45 50 55						
GCT CCC AGC CTG AGG CCT GCC CCA CCG CCC ATC AGT GGA GGT GGC TAT	3356					
Ala Pro Ser Leu Arg Pro Ala Pro Pro Pro Ile Ser Gly Gly Gly Tyr						
60 65 70						
CGG GCT CGT CCA GCC AAA GCA GCT GCC ACT CAA AAG AAA GTA GAA AGA	3404					
Arg Ala Arg Pro Ala Lys Ala Ala Ala Thr Glu Lys Lys Val Glu Arg						
75 80 85						
AAA GCC CCT GAT GCT GGA GGC TGT CTT CAC GCT GAC CCA GAC CTG	3449					
Lys Ala Pro Asp Ala Gly Gly Cys Leu His Ala Asp Pro Asp Leu						
90 95 100						
GTGGGTGCAC TGATGTTTTCT TCAGTGGTG GCTCTCTCAT GCAGAGAAAAG CCTGTAGTCA	3509					
TGGCAGTCTG CTAATGTTTC ACTGACCCAC ATTACCATCA CTGTTATTTT GTTTGTTTTAT	3569					
TTTGGAAATA AAATTCAAAA CATAAACATA TTGGGCCTTT GGTTTAAGCT TTCTTTCTTG	3629					
TTTTCTTTGG TCTGGGCCCA AAATTTCAAA TTAGGATATG TGGGTGCCAC CTTTCCATTT	3689					
GTATTTTGCC ACTGCCITTT TTTAGTTGGT AAAATTTTCA TAGCCCAATT ATATTTTTTC	3749					
TGGGTAAGT AATATTTTAA ATCTCTATGA GAGTATGATG ATGACTTTTG AATTTCTGGT	3809					
CTTACAGAAA ACCAAATAAT AAATTTTTAT GTTGGCTAAT CGTATCGCTG AATTTTCCTA	3869					
TGTGCTATTT TAACAAATGT CCATGACCCA AATCCTTCAT CTAATGCCCTG CTATTTTCTT	3929					
TGTTTTTAAG GGG GTG TTG TGT CCT ACA GGA TGT CAG TTG CAA GAG GCT	3977					
Gly Val Leu Cys Pro Thr Gly Cys Glu Leu Glu Glu Ala						
105 110 115						
TTG CTA CAA CAG GAA AGG CCA ATC AGA AAT AGT GTT GAT GAG TTA AAT	4025					
Leu Leu Glu Glu Glu Arg Pro Ile Arg Asn Ser Val Asp Glu Leu Asn						
120 125 130						
AAC AAT GTG GAA GCT GTT TCC CAG ACC TCC TCT TCT TCC TTT CAG TAC	4073					
Asn Asn Val Glu Ala Val Ser Glu Thr Ser Ser Ser Ser Phe Glu Tyr						
135 140 145						
ATG TAT TTG CTG AAA GAC CTG TGG CAA AAG AGG CAG AAG CAA GTA AAA G	4122					
Met Tyr Leu Leu Lys Asp Leu Trp Glu Lys Arg Glu Lys Glu Val Lys						
150 155 160						
GTAGATATCC TTGTGCTTTC CATTGATTT TCAGCTATAA AATTGGAACC GTTAGACTGC	4182					
CACGAGAAAT CATGGTTGTG AGAAGATTAA CATTCTGGG TTAGTGAATA GCATTTCATAC	4242					
GCTTTTGGGC ACCTTCCCTT GCAACTTGCC AGATAAGCAC TATTCAAGCTC TTATTCCCAG	4302					
TCTGACATCA GCAAGTGTGA TTTTCTATGA AAAATTCTAC TATGACTCCT TATTTTAAAT	4362					
ATACAAAGAAA CTGTGACTC AGAAGATAAT ATTTACAGAG TGGAAAAAAA CCCCTAGCAT	4422					
TTATAATTTT AACATTTGAG OTTTTGAATG AGAGAATTAT CCATAATATA TTCAATTGTG	4482					
TTGTGGATAA TGACACCTAA CCTGTGAATC TTGAGGTGAG AATGTTGAGT GCTGTTGACT	4542					
TGGTGGTCAG GAAACAOGTA GTGCGTGAGC CTGGCACAGG CATCTCAGTG AGTAGCATAC	4602					
CCACAGTTGG AAATTTTTCA AAGAAATCAA AOGAATCATG ACATCTTATA AATTTCAAAG	4662					
TTCTGCTATA CTTATGTGAA ATGGATAAAT AAATCAAGCA TATCCACTCT GTAAAGATTGA	4722					
ACTTCTCAGA TGGAAAGACCC CAATACTGCT TTCTCCTCTT TTCCCTCACC AAAAGAAATAA	4782					
ACAACCTATT TCATTTATTA CTGGACACAA TCTTTAGCGT ATACCTATGG TAAATTACTA	4842					

-continued

GTATGGTGGT TAGGATTTAT GTTAATTTGT ATATGTCATG CGCCAAATCA TTTCCACTAA	4902
ATATGACTAT ATATCATAAC TGCTTGGTGA TAGCTCAGTG TTTAATAGTT TATTCTCAGA	4962
AAATCAAAAT TGTATAGTTA AATACATTAG TTTTATGAGG CAAAAATGCT AACTATTTCT	5022
ACATAATTTT ATTTTTCAG AT AAT GAA AAT GTA GTC AAT GAG TAC TCC	5071
<div> <div>Asp Asn Glu Asn Val Val Asn Glu Tyr Ser</div> <div>165 170</div> </div>	
TCA GAA CTG GAA AAG CAC CAA TTA TAT ATA GAT GAG ACT GTG AAT AGC	5119
Ser Glu Leu Glu Lys His Glu Leu Tyr Ile Asp Glu Thr Val Asn Ser	
175 180 185	
AAT ATC CCA ACT AAC CTT COT GTG CTT COT TCA ATC CTG GAA AAC CTG	5167
Asn Ile Pro Thr Asn Lys Arg Val Leu Arg Ser Ile Leu Glu Asn Leu	
190 195 200 205	
AGA AGC AAA ATA CAA AAG TTA GAA TCT GAT GTC TCA OCT CAA ATG GAA	5215
Arg Ser Lys Ile Glu Lys Leu Glu Ser Asp Val Ser Ala Glu Met Glu	
210 215 220	
TAT TOT CGC ACC CCA TGC ACT GTC AGT TGC AAT ATT CCT GTG GTG TCT	5263
Tyr Cys Arg Thr Pro Cys Thr Val Ser Cys Asn Ile Pro Val Val Ser	
225 230 235	
GGC AAA G GTAACGATT CATAAACATA TTTTATGAGA GTTCCAGAAAG AACTCACACA	5320
Gly Lys	
CCAAAAATAA GAGAACAAACA ACAACAACAA AAATGCTAAG TGGATTTTCC CAACAGATCA	5380
TAATGACATT ACAGTACATC ATAAAAATAT CCTTAGCCAG TTGTGTTTTG GACTGGCCTG	5440
GTGCATTTGC TGGTTTTGAT GAGCAGGATG GGGCAGAGGT AGTCCCAAGG GTGGCTGATG	5500
TGTGCATCTG CGTACTGGCT TGAACAGATG GCAQAACCAC AGATAGATGT AGAAGTTTCT	5560
CCATTTTGTG TGTCTGGGA GCTCATGGAT ATTCCAAGAC AAAAAAGGTG GAGAAAGAGCT	5620
TTGTTTATCC TCTTAGCAGA TAAACGTCTT CAAACTGGG TTGGACTTAC TAAAGTAAAA	5680
TGAAAAATCTA ATATTTGTTA TATTATTTTC AAAAGTCTAT AATAACACAC TCCTTAGTAA	5740
CTTATGTAAT GTTATTTTAA AGAATTGGTG ACTAAATACA AAGTAATTAT GTCATAAACCC	5800
CCTGAACATA ATGTTGTCTT ACATTTGCAG AA TGT GAG GAA ATT ATC AGG AAA	5853
<div> <div>Glu Cys Glu Glu Ile Ile Arg Lys</div> <div>240 245</div> </div>	
GGA GGT GAA ACA TCT GAA ATG TAT CTC ATT CAA CCT GAC AGT TCT GTC	5901
Gly Gly Glu Thr Ser Glu Met Tyr Leu Ile Glu Pro Asp Ser Ser Val	
250 255 260	
AAA CCG TAT AGA GTA TAC TGT GAC ATG AAT ACA GAA AAT GGA G	5944
Lys Pro Tyr Arg Val Tyr Cys Asp Met Asn Thr Glu Asn Gly	
265 270 275	
GTAAAGCTTT GACAGTTGTT GACCTGTTGA TCTGTAATTA TTTGGATACC GTAAAAATGCC	6004
AGGAAACAAAG GCCAGGTGTG GTGGCTCATA CCTGTAATTC CAGCACCTTG GGAGGCCAAA	6064
GTGGGCTGAT AGCTTGAGCC TAGGAGTTTG AAAGTACCTT GGGCAACATA ATGAGACCTT	6124
AACTCTACAA AAAAAAAAAA AATACCAAAA AAAAAAAAAA AATCAGCTGT GTTGGTAGTA	6184
TGTGCCTGTA GTCCCAAGCTA TCCAAGAGGC TGAATAGGGA GATCACCTGA GCCCACAACC	6244
TGGAGTCTTG ATCATGCTAC TGAAGTGTAG CCTGGGCAAC AGAGGATAGT GAGATCCTGT	6304
CTCAAAAAAA AAAATTAATT AAAAAAGCCAG GAAACAAGAC TTAGCTCTAA CATCTAACAT	6364
AGCTGACAAA GGAGTAATTT GATGTGGAAT TCAACCTGAT ATTTAAAAAT TATAAAATAT	6424
CTATAATTCA CAATTTGGGG TAAAGATAAAG CACTTGCAAT TTCCAAAAGT TTTACAAATT	6484
TACCTCTCAT ATTTATTTCC TTATTGTGTC TATTTTAGAG CACCAAATAT ATACTAAATG	6544
GAATGGAACG GGGATTGAGA TATTATTTTC AAAAGTGACAT TATTTGCTGT TGGTTAATAT	6604

-continued

ATGCTCTTTT	TGTTTCTGTC	AACCAAAAG	GA	TGG	ACA	GTG	ATT	CAG	AAC	CGT	6653
Gly Trp Thr Val Ile Glu Asn Arg 280 285											
CAA	GAC	GGT	AGT	GTT	GAC	TTT	GGC	AAG	AAA	TGG	6703
Gln	Asp	Gly	Ser	Val	Asp	Phe	Gly	Arg	Lys	Trp	
290 295 300											
GGA	TTT	GGA	AAT	GTT	GCA	ACC	AAC	ACA	GAT	GGG	6751
Gly	Phe	Gly	Asn	Val	Ala	Thr	Asn	Thr	Asp	Gly	
305 310 315											
CTA	CCA	G	GTAACGAACA	GGCATGCAAA	ATAAAATCAT	TCTATTTGAA	ATGGGATTTT	6808			
Leu	Pro										
TTTTAATTAA	AAAACATTCA	TTGTTGGAAG	CCTGTTTTAG	GCAGTTAAAG	GGAGTTTCCT	6868					
GACAAAAATG	TGGAAGCTAA	AGATAAGGGA	AGAAAAGGCAG	TTTTTAGTTT	CCCAAAATTT	6928					
TATTTTTGGT	GAGAGATTTT	ATTTTGTTTT	TCTTTTAG	GT	GAA	TAT	TGG	CTT	6980		
Gly Glu Tyr Trp Leu 320											
GGA	AAT	GAT	AAA	ATT	AGC	CAG	CTT	ACC	AAG	ATG	7028
Gly	Asn	Asp	Lys	Ile	Ser	Glu	Leu	Thr	Arg	Met	
325 330 335 340											
TTG	ATA	GAA	ATG	GAG	GAC	TGG	AAA	GGA	GAC	AAA	7076
Leu	Ile	Glu	Met	Glu	Asp	Trp	Lys	Gly	Asp	Lys	
345 350 355											
GGA	GGA	TTC	ACT	GTA	CAG	AAT	GAA	GCC	AAC	AAA	7124
Gly	Gly	Phe	Thr	Val	Glu	Asn	Glu	Ala	Asn	Lys	
360 365 370											
AAC	AAA	TAC	AGA	GGA	ACA	GCC	GGT	AAT	GCC	CTC	7172
Asn	Lys	Tyr	Arg	Gly	Thr	Ala	Gly	Asn	Ala	Leu	
375 380 385											
CAG	CTG	ATG	GGA	GAA	AAC	AAG	ACC	ATG	ACC	ATT	7220
Gln	Leu	Met	Gly	Glu	Asn	Arg	Thr	Met	Thr	Ile	
390 395 400											
TTC	AGC	ACG	TAT	GAC	AGA	GAC	AAT	GAC	GGC	TG	7262
Phe	Ser	Thr	Tyr	Asp	Arg	Asp	Asn	Asp	Gly	Trp	
405 410 415											
CAC	TCTTTTC	TCCTGCTTTA	AAAATCACAC	TAATATCATT	ACTCAGAATC	ATTAACAATA	7322				
TTTTAATAAG	CTACCACTTC	CTGGGCACCT	ACTGTCAGCC	ACTGTCCTAA	GCTCTTTTATG	7382					
CATCACTCGA	AAOCATTTC	ACTATAAAGT	AGACATTCTT	ATTCTCATTT	TACAGATGAG	7442					
ATTTAAGAGAG	ATTACGTGAT	TTGTCCAATG	TCACACAACCT	ACCCAGAGAT	AAAACCTAGAA	7502					
TTTGAGCACA	GTTACTTTCT	GAATAATGAG	CATTTAAGATA	AATACCTATA	TCTCTATATT	7562					
CTAAAGTGTO	TGTGAAAACT	TTCATTTTCA	TTTCCAAGGT	TCTCTGATAC	TAAAGGTTGT	7622					
AAAAGCTATT	ATTCCAGTAT	AAAAGTAACAA	ACACAGTCCC	TAGATGGATT	GCCACAAAAG	7682					
CCCAATTATC	TCTCTTTCTT	GCTATAAGGC	ACAAGGAGTC	TTTGGTGAT	TAGTGCTACT	7742					
CTATGTATAG	CACCCAAAAG	AAAGACTACT	GTGCACACGA	GTGTAGCAGT	CTTTTATGAG	7802					
TAATCTGCAA	AACGTAACTT	GACCACCCTA	GTTCTGTTTC	TAATAACGCC	AAACACATTT	7862					
TCITTCAAG	G	TTA	ACA	TCA	GAT	CCC	AGA	AAA	CAG	TGT	7910
Leu Thr Ser Asp Pro Arg Lys Glu Cys Ser Lys Glu Asp 420 425											
GGT	GGT	GGA	TGG	TGG	TAT	AAT	AGA	TGT	CAT	GCA	7958
Gly	Gly	Gly	Trp	Trp	Tyr	Asn	Arg	Cys	His	Ala	
430 435 440											
AGA	TAC	TAC	TGG	GGT	GGA	CAG	TAC	ACC	TGG	GAC	8006
Arg	Tyr	Tyr	Trp	Gly	Gly	Glu	Tyr	Thr	Trp	Asp	
Met Ala Lys His Gly											

-continued

445	450	455	460	
ACA GAT GAT GGT GTA GTA TGG ATG AAT TGG AAG GGG TCA TGG TAC TCA				8054
Thr Asp Asp Gly Val Val Trp Met Asn Trp Lys Gly Ser Trp Tyr Ser				
	465	470	475	
ATG AAG AAG ATG AGT ATG AAG ATC AAG CCC TTC TTC CCA CAG CAA TAGTCCCC				8109
Met Arg Lys Met Ser Met Lys Ile Arg Pro Phe Phe Pro Gln Gln				
	480	485	490	
TACGTAGATT TTGCTCTTC TGTATGTGAC AACATTTTTG TACATTATGT TATTGGAATT				8169
TTCTTTCATA CATTATATTC CTCTAAAACT CTCAAGCAGA CGTGAGTGTG ACTTTTTGAA				8229
AAAAAGTATAG GATAAATTAC ATTAAAAATAG CACATGATTT TCTTTTGTTT TCTTCATTTT				8289
TCTTGCTCAC CCAAGAAAGTA ACAAAAAGTAT AGTTTTGACA GAGTTGGTGT TCATAATTTT				8349
AGTTCTAGTT GATTGCGAGA ATTTTCAAAT AAGGAAGAGG GGTCTTTTAT CCTTGTCTGA				8409
GGAAAAACCAT GACCGAAAGG AAAAACTGAT GTTTAAAAAGT CCACTTTTAA AACTATATTT				8469
ATTTATGTAG GATCTGTCAA AGAAAACTTC CAAAAAGATT TATTAATTAA ACCAGACTCT				8529
GTTCGAATAA GTTAATGTTT TCTTGTTTTG TAATCCACAC ATTCAATGAG TTAGGCTTTG				8589
CACTTGTAAAG GAAGGAGAAAG CTTTCACAAC CTCAAATAGC TAATAAACCG GTCTTGAATA				8649
TTTGAAGATT TAAAATCTGA CTCTAGGACG GGCACGGTGG CTCACGACTA TAATCCCAAC				8709
ACTTTGGGAG GCTGAGGCGG GCGGTCACAA GGTCAAGGAGT TCAAGACCAO CCTGACCAAT				8769
ATGGTGAAAC CCCATCTCTA CTAAAAATAC AAAAAATTAGC CAGGCGTGGT GGCAGGTGCC				8829
TGTAGGTCCC AGCTAGCCTG TGAGGTGGAG ATTGCATTGA GCCAAGATC				8878

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 491 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Lys	Arg	Met	Val	Ser	Trp	Ser	Phe	His	Lys	Leu	Lys	Thr	Met	Lys	
1				5					10					15		
His	Leu	Leu	Leu	Leu	Leu	Leu	Cys	Val	Phe	Leu	Val	Lys	Ser	Gln	Gly	
			20					25					30			
Val	Asn	Asp	Asn	Glu	Glu	Gly	Phe	Phe	Ser	Ala	Arg	Gly	His	Arg	Pro	
		35					40					45				
Leu	Asp	Lys	Lys	Arg	Glu	Glu	Ala	Pro	Ser	Leu	Arg	Pro	Ala	Pro	Pro	
	50				55						60					
Pro	Ile	Ser	Gly	Gly	Gly	Tyr	Arg	Ala	Arg	Pro	Ala	Lys	Ala	Ala	Ala	
65			70				75							80		
Thr	Gln	Lys	Lys	Val	Glu	Arg	Lys	Ala	Pro	Asp	Ala	Gly	Gly	Cys	Leu	
			85				90							95		
His	Ala	Asp	Pro	Asp	Leu	Gly	Val	Leu	Cys	Pro	Thr	Gly	Cys	Gln	Leu	
		100					105						110			
Gln	Glu	Ala	Leu	Leu	Gln	Gln	Glu	Arg	Pro	Ile	Arg	Asn	Ser	Val	Asp	
		115				120					125					
Glu	Leu	Asn	Asn	Asn	Val	Glu	Ala	Val	Ser	Gln	Thr	Ser	Ser	Ser	Ser	
	130				135						140					
Phe	Gln	Tyr	Met	Tyr	Leu	Leu	Lys	Asp	Leu	Trp	Gln	Lys	Arg	Gln	Lys	
145				150				155						160		
Gln	Val	Lys	Asp	Asn	Glu	Asn	Val	Val	Asn	Glu	Tyr	Ser	Ser	Glu	Leu	
			165					170						175		

-continued

Glu	Lys	His	Gln	Leu	Tyr	Ile	Asp	Glu	Thr	Val	Asn	Ser	Asn	Ile	Pro	180	185	190
Thr	Asn	Leu	Arg	Val	Leu	Arg	Ser	Ile	Leu	Glu	Asn	Leu	Arg	Ser	Lys	195	200	205
Ile	Gln	Lys	Leu	Gln	Ser	Asp	Val	Ser	Ala	Gln	Met	Glu	Tyr	Cys	Arg	210	215	220
Thr	Pro	Cys	Thr	Val	Ser	Cys	Asn	Ile	Pro	Val	Val	Ser	Gly	Lys	Glu	225	230	235
Cys	Glu	Glu	Ile	Ile	Arg	Lys	Gly	Gly	Glu	Thr	Ser	Glu	Met	Tyr	Leu	245	250	255
Ile	Gln	Pro	Asp	Ser	Ser	Val	Lys	Pro	Tyr	Arg	Val	Tyr	Cys	Asp	Met	260	265	270
Asn	Thr	Glu	Asn	Gly	Gly	Trp	Thr	Val	Ile	Gln	Asn	Arg	Gln	Asp	Gly	275	280	285
Ser	Val	Asp	Phe	Gly	Arg	Lys	Trp	Asp	Pro	Tyr	Lys	Gln	Gly	Phe	Gly	290	295	300
Asn	Val	Ala	Thr	Asn	Thr	Asp	Gly	Lys	Asn	Tyr	Cys	Gly	Leu	Pro	Gly	305	310	315
Glu	Tyr	Trp	Leu	Gly	Asn	Asp	Lys	Ile	Ser	Gln	Leu	Thr	Arg	Met	Gly	325	330	335
Pro	Thr	Glu	Leu	Leu	Ile	Glu	Met	Glu	Asp	Trp	Lys	Gly	Asp	Lys	Val	340	345	350
Lys	Ala	His	Tyr	Gly	Gly	Phe	Thr	Val	Gln	Asn	Glu	Ala	Asn	Lys	Tyr	355	360	365
Gln	Ile	Ser	Val	Asn	Lys	Tyr	Arg	Gly	Thr	Ala	Gly	Asn	Ala	Leu	Met	370	375	380
Asp	Gly	Ala	Ser	Gln	Leu	Met	Gly	Glu	Asn	Arg	Thr	Met	Thr	Ile	His	385	390	395
Asn	Gly	Met	Phe	Phe	Ser	Thr	Tyr	Asp	Arg	Asp	Asn	Asp	Gly	Trp	Leu	405	410	415
Thr	Ser	Asp	Pro	Arg	Lys	Gln	Cys	Ser	Lys	Glu	Asp	Gly	Gly	Gly	Trp	420	425	430
Trp	Tyr	Asn	Arg	Cys	His	Ala	Ala	Asn	Pro	Asn	Gly	Arg	Tyr	Tyr	Trp	435	440	445
Gly	Gly	Gln	Tyr	Thr	Trp	Asp	Met	Ala	Lys	His	Gly	Thr	Asp	Asp	Gly	450	455	460
Val	Val	Trp	Met	Asn	Trp	Lys	Gly	Ser	Trp	Tyr	Ser	Met	Arg	Lys	Met	465	470	475
Ser	Met	Lys	Ile	Arg	Pro	Phe	Phe	Pro	Gln	Gln						485	490	

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10564 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

- (B) CLONE: human fibrinogen gamma chain

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(1799..1876, 1973..2017, 2207..2390, 2510
... 2603, 4211..4341, 4645..4778, 5758..5942, 7426
... 7703, 9342..9571)

-continued

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTACACACTT	CTTGAAGGCA	AAGGCAATGC	TGAAGTCACC	TTTCATGTTT	AAATCATATT	60
AAAAAGTTAG	CAAGATGTAA	TTATCAGTGT	ACTATGTAAA	TCTTTGTGAA	TGATCAATAA	120
TTACATATTT	TCATTATATA	TATTTTAAGT	GATAATATTT	ATATACATTC	AACATTCTAA	180
ATATAGAAAAG	TTTACAGAGA	AAAATAAAGC	CTTTTTTTTC	AATCCTGTCC	TCCACCTCTG	240
CATCCCATTG	TTCTTCACAG	AAGCAACTGA	TTCAAAGTCAT	TACATAGTTA	TTGAAGTGTTA	300
ACTACAACCTA	TGTTAAGTAC	AGCTATATAT	GTTAGATGCC	GTAGCCACAG	AAATCAGTTT	360
ACAATCTAAT	GCAGTGGATA	CAGCATGTAT	ACATATAATA	TAAAGTTGCT	ACAAATGCTA	420
TCTGAAGGTAG	AGCTGTTTGA	AAGAATACTA	ATACTTAAAT	GTTTAATTCA	ACTGACTTGA	480
TTGACAACTG	ATTAGCTGAG	TGGAAAAGAT	GGATGAGAAA	GATTGTGAGA	CTTAATTGGC	540
TGGTGGTATG	GTGATATGAT	TGACAATAAC	TGCTAAAGTCA	GAGAGGGGATA	TATTAAGGAG	600
GAGAAAGAAAA	GCAACAAATC	TGGTTTTGAT	GTGTTCACTT	TGTTATAATT	ATTGATTATT	660
TACTGAATAT	GAATATTTAT	CTTTGTTTTT	GAGTCAAT	AA	ATATACCTTT	720
AAITAAAAGTA	TTAGTATTTT	TTTCAAACTG	GAGGCATTTT	TCCCACCTAAC	ATATTTTCATC	780
AAAACCTTATA	ATAAGCTTGG	TTCCAAGGGA	AGAAATGAGG	GATAACCCAAA	AATAGAGACA	840
TTAATAATAG	TGTAACGCCC	AGTGATAAAT	CTCAATAGGC	AGTGATGACA	GACATGTTTT	900
CCCAAAACACA	AGGATGCTGT	AAAGGCCAAA	CAGAAATGAT	GGCCCTCCC	CAGCACCTCA	960
TTTTGCCCCI	TCCTTCAGCT	ATGCCCTCTAC	TCTCCTTTAG	ATACAAAGGGA	GGTGGATTTT	1020
TCTCTTCTCT	GAGATAGCTT	GATGGAACCA	CAGGAACAAT	GAAAGTGGGCT	CCTGGCTCTT	1080
TTCTCTGTGG	CAGATGGGGT	GCCATGCCCA	CCTTCAGACA	AAAGGAAAGAT	TGAGCTCAAA	1140
AGCTCCCTGA	GAAAGTGAGAG	CCTATGAACA	TGGTTGACAC	AGAGGGGACAG	GAATGTATTT	1200
CCAGGGTCAT	TCATTCCCTGG	GAATAGTGAA	CTGGGACATG	GGGGAAAGTCA	GTCTCCTCCT	1260
GCCACAGCCA	CAGATTAAAA	ATAATAATGT	TAACTGATCC	CTAGGCTAAA	ATAATAAGTG	1320
TAACTGATCC	CTAAAGCTAAG	AAAATTCTTT	TGGTAATTCA	GGTGATGGCA	GCAAGACCCA	1380
TCTTAAAGAT	AGACTAGGTT	TGCTTAAGTT	GAGGTCAAT	CTGTTTGCTC	TCAAGCATGT	1440
ACTGGAAGAA	GTTCATCACC	ACAACCTCCA	GGACTGCCCT	CCTCCTCACA	GCAATGGATA	1500
ATGCTTCACT	AGCCTTTTGA	GATAATTTTG	GATCAGAGAA	AAAACCTTGA	GCTGGGCCAA	1560
AAAAGGAGAG	CTTCAACCTG	TGTGCAAAAT	CTGGGAACCT	GACAATATAAG	GTGAGGGGCC	1620
AGGATGAGGA	AAAAAGGAAC	GGAAAGACCT	GGCCACCCCT	CTGGTAAAGGA	GGCCCCGTGA	1680
TCAAGCTCCAG	CCATTTGCAAG	TCCTGGCTAT	CCCAAGAGCT	TACATAAAGG	GACAATTGGA	1740
GGCTGAGAGG	TGACAGTGCT	GACACTACAA	GGCTCGGAGC	TCCGGGCACT	CAGACATC	1798
ATG AGT TGG TCC TTT CAC CCC CGG AAT TTA ATT CTC TAC TTC TAT GCT	1846					
Met Ser Trp Ser Leu His Pro Arg Asn Leu Ile Leu Tyr Phe Tyr Ala						
1 5 10 15						
CTT TTA TTT CTC TCT TCA ACA TGT GTA GCA GTAAAGTGTGC TCTTCACAAA	1896					
Leu Leu Phe Leu Ser Ser Thr Cys Val Ala						
20 25						
ACGTTGTTTA AAATGGAAAAG CTGGAAAATA AAACAGATAA TAAACTAGTG AAATTTTCGT	1956					
ATTTTTCTC TTTTAG TAT GTT GCT ACC AGA GAC AAC TGC TGC ATC TTA	2005					
Tyr Val Ala Thr Arg Asp Asn Cys Cys Ile Leu						
30 35						
GAT GAA AGA TTC GTAAAGTAGTT TTTATGTTTC TCCCTTTGTG TGTGAAGTGG	2057					
Asp Glu Arg Phe						

-continued

40

AGAGGGGOCAG AGGAATAGAA ATAATTCCT CATAAATATC ATCTGGCACT TGTAACTTTT 2117
 TAAAAACATA GTCTAGGTTT TACCTATTTT TCTTAATAGA TTTTAAAGAGT AGCATCTGTC 2177
 TACATTTTTTA ATCACTGTTA TATTTTCAO GGT AGT TAT TGT CCA ACT ACC TGT 2230
 Gly Ser Tyr Cys Pro Thr Thr Cys
 45
 GGC ATT GCA GAT TTC CTG TCT ACT TAT CAA ACC AAA GTA GAC AAG GAT 2278
 Gly Ile Ala Asp Phe Leu Ser Thr Tyr Glu Thr Lys Val Asp Lys Asp
 50 55 60 65
 CTA CAG TCT TTG GAA GAC ATC TTA CAT CAA GTT GAA AAC AAA ACA TCA 2326
 Leu Glu Ser Leu Glu Asp Ile Leu His Glu Val Glu Asn Lys Thr Ser
 70 75 80
 GAA GTC AAA CAG CTG ATA AAA GCA ATC CAA CTC ACT TAT AAT CCT GAT 2374
 Glu Val Lys Glu Leu Ile Lys Ala Ile Glu Leu Thr Tyr Asn Pro Asp
 85 90 95
 GAA TCA TCA AAA CCA A GTGAGAAAAT AAAGACTACT GACCAAAAAA 2420
 Glu Ser Ser Lys Pro
 100
 TAATAATAAT AATCTGTGAA GTTCTTTTGC TGTGTTTTAA GTTGTTCAT TTGCTTAAAG 2480
 ATTTTATGT CTCTGATCCT ATATTACAG AT ATG ATA GAC GCT GCT ACT TTG 2532
 Asn Met Ile Asp Ala Ala Thr Leu
 105 110
 AAG TCC AGG ATA ATG TTA GAA GAA ATT ATG AAA TAT GAA GCA TCG ATT 2580
 Lys Ser Arg Ile Met Leu Glu Glu Ile Met Lys Tyr Glu Ala Ser Ile
 115 120 125
 TTA ACA CAT GAC TCA AGT ATT CG GTAAGGATTT TTGTTTTAAT TTGCTCTOCA 2633
 Leu Thr His Asp Ser Ser Ile Arg
 130
 AGACTGATTT AGTTTTTATT TAATATTCTA TACTTGAAGTG AAAGTAATTT TTAATGTGTT 2693
 TTCCCCATTT ATAATATCCC AGTOACATTA TGCCGTGATTA TGTTOAGCAT AGTAAGAGATA 2753
 GAAATTTTTTA GTGCAATATA AATTATACTG GOTTATAATT GCTTATTAAT AATCACATTO 2813
 AAGAAAAGATG TTCTAGATGT CTTCAAAATGC TAOTTTGACC ATATTTATCA AAAATTTTTTT 2873
 CCCCATCCCC CATTTATCTT ACAACATAAA ATCAATCTCA TAGGAATTTG GGTGTTGAAA 2933
 ATAAAAATCCT CTTTATAAAA ATGCTGACAA ATTGGTGTTT AAAAAAATTA GCAAAGCAGAG 2993
 GCATAGTAAO GATTTTGGCT CCTAAAGTAA ATTATATTGA ATGTGGAGCA GGAAGAAACA 3053
 TOTCTTGAAG GACTAAGTGT GGCAAAATATT GCAAAAGCTCA TATTGATCAT TGCAGAATGA 3113
 ACCTGCATAO TCTCTTCCCT TCAATTTGAA GTGAATGTCT CTGTTAAAGC TTCTCAAGGA 3173
 CTCATAAACT TTCTGAACAT AAGGTCTCAG ATACAGTTTT AATATTTTTT CCAATTTTTT 3233
 TTTTCTGAAT TTTTCTCAAA GCAGCTTGAO AAATTOAGAT AAATAGTAGC TAGGAGAGAG 3293
 TGGCCCAAGG AAGATTTCTC CTCTTTTTGC TATCAAGAGG CCCTTGTTAT TATTGTTATT 3353
 ATTATTACTT GCATTATTAT TOTCCATCAT TGAAGTTGAA GGAGGTTATT GTACAGAAAT 3413
 TGCCTAAGAC AAGGTAGAGG GAAAACGTGG ACAAATAGTT TGTCTACCCT TTTTACTTC 3473
 AAGAAAAGAA CGTTTATGC ATTGTAGACA GTTTTCTATC ATTTTGGAT ATTTGCAAGC 3533
 CACCCTGTAA GTAACACAA AAGGAAGGTT TTTACTTCCC CCAATCCATT CCCAAAAGCTA 3593
 TOTAACCAGA AGCATTAAAG AAGAAAAGGG AAGTATCTGT TTTTTTATT TACATACAAT 3653
 AACGTTCCAG ATCATGTCCC TGTGTAAATT ATATTTTGA TGAAGCTTA TATGTATAGC 3713
 CTCAGTAGAT CCACAAGTGA AAGGTATACT CCTTCAGCAC ATGTGAATTA CTGAAGTAG 3773

-continued

CTTTTCCTGC	TTCTAAAGCA	TCAGGGGGTG	TTCCTATTAA	CCAATCTCOC	CACTCTTGCA	3833
GGTTGCTATC	TGCTGTCCCT	TATGCATAAA	GTAAAAAGCA	AAATGTCAAT	GACATTTGCT	3893
TATTGACAAO	GACTTTGTTA	TTTGTGTTGG	GAGTTGAGAC	AATATGCCCC	ATTCTAAGTA	3953
AAAAGATTCA	GGTCCACATT	GTATTCCTGT	TTTAATTGAT	TTTTTGATTT	GTITTTCTTT	4013
TTCAAAAAAGT	TTATAATTTT	AATTCATGTT	AATTTAGTAA	TATAATTTTA	CATTTTCCTC	4073
AAGAATGGAA	TAATTTATCA	GAAAGCACTT	CTTAAGAAAA	TACTTAGCAG	TTTCCAAAQA	4133
AAATATAAAA	TTACTCTTCT	GAAAGGAATA	CTTATTTTTG	TCTTCTTATT	TTTGTATCT	4193
TATGTTTCTG	TTTGTAG A	TAT TTG CAG	GAA ATA TAT	AAT TCA AAT	AAT CAA	4244
Tyr Leu Glu Glu Ile Tyr Asn Ser Asn Asn Glu 135 140 145						
AAG ATT GTT AAC CTG	AAA GAG AAG GTA	GCC CAG CTT GAA	GCA CAG TGC			4292
Lys Ile Val Asn Leu	Lys Glu Lys Val	Ala Glu Leu Glu	Ala Glu Cys			
150 155 160						
CAG GAA CCT TGC	AAA GAC ACO GTG	CAA ATC CAT GAT	ATC ACT GGG AAA	G		4341
Glu Glu Pro Cys	Lys Asp Thr Val	Glu Ile His Asp	Ile Thr Gly Lys			
165 170 175						
GTAACTGATG	AAGGTTATAT	TGGGATTAGG	TTCATCAAAO	TAAATAATGT	AAAGGAGAAA	4401
GTATGTACTG	GAAAGTATAG	GAATAGTTTA	GAAAGTGOGT	ACCCATTAAO	TCTAAGAATT	4461
TCAOTTGTCT	AGACCTTTCT	TOAATAOCTA	AAAAAAACAG	TTTAAAAOGA	ATGCTGATGT	4521
GAAAAGTAAO	AAAAATTATC	TTGGAAAAATG	AATAGTTTAC	TACATGTTAA	AAOCTATTTT	4581
TCAAGOGCTOO	CACAOTCTTA	CCTOCATTTT	AAACCACAGT	AAAAOTCGAT	TCTCCTTCTC	4641
TAG AT TGT	CAA GAC ATT	GCC AAT AAG	GGA GCT AAA	CAG AOC	GGG CTT	4688
Asp Cys	Glu Asp Ile	Ala Asn Lys	Gly Ala Lys	Glu Ser	Gly Leu	
180 185 190						
TAC TTT ATT	AAA CCT CTG	AAA OCT AAC	CAG CAA TTC	TTA GTC	TAC TGT	4736
Tyr Phe Ile	Lys Pro Leu	Lys Ala Asn	Glu Glu Glu	Phe Leu	Val Tyr Cys	
195 200 205						
GAA ATC GAT	GGG TCT GGA	AAT GGA TGG	ACT GTG T--	CAG AAG		4778
Glu Ile Asp	Gly Ser Gly	Asn Gly Trp	Thr Val Phe	Glu Lys		
210 215 220						
GTAAATTTTTT	CCCCACCATG	TGTATTTAAT	AAATTCCTAC	ATTGTTTCTG	CCATATGQCA	4838
GATACTTTTTT	TAAOCACCTT	GTGAACCOTA	GCTCATTTAA	TCCTTQCAAT	AGCCCTAAQA	4898
OGAAOGTACT	TCTGTTACTC	CTATTTACAG	AAAAOGAAAC	TGAOGCACAC	AAOGTTAAAT	4958
AACTTGCCCA	AGACCACATA	ACTAATAAOC	AACAAGATCA	GCATTTGAAC	CTAGGCAQTA	5018
TAGTTTCAQA	OTTTGTGACT	TOACTCTATA	TTGTACTGOC	ACTGACTTTO	TAGATTGATO	5078
GTGGCACATA	ATCATAOTAC	CACAGTOACA	AATAAAAAQA	AGGAAACTCT	TTTGTCAOGT	5138
AOGTCAAGAC	CTGAGGTTTC	CCATCACAAO	ATGAGGAAOC	CCAAACACCAC	CCCCCACCAC	5198
CCCACCACCA	TCACCACCCT	TTACACACACC	AGAOGATACA	CTTGGGCTGC	TCCAAOACAA	5258
OGAACCTOTO	TTGCATCTOC	CACTTOCTOA	TACCCACTAG	GAATCTTGGC	TCCTTTACTT	5318
TCTGTTTACC	TCCCACCACCT	OTTATAACTO	TTTCTACAGG	GGGCGCTCAO	AGGGAATGAA	5378
TGGTGGAAOC	ATTAGTTGCC	AGACACCOAT	TGAOGCAATGG	GTTCATCAT	AAOTGTAAQA	5438
ATCAQTAATA	TCCAQCTAQA	GTCTGAAAGT	CGTCTAGGTG	TCTTTTTAAT	ATTACCACTC	5498
ATTTAGAAAT	TATGATGTGC	CAGAAACCCCT	CTTAAGTATT	TCTCTTATAT	TCTCTCTCAT	5558
GATCCTTQCA	OCAACCCTAA	GAAGTAACCA	TCATTTTTTCC	TATTTGATAC	ATGAGGAAAC	5618
TGAOGTAOCT	TGGCCAAGAT	CACTTAOTTO	OGAOTTOATA	GAACCAOTGC	TCTGTATTTT	5678
TGACAAAAATG	TTGACAGCAT	TCTCTTTACA	TGCATTGATA	GTCTATTTTC	TCCTTTTGCT	5738

-continued

CTTGCAAATG TGTAATTAG AGA CTT GAT GGC AGT GTA GAT TTC AAG AAA AAC	5790
Arg Leu Asp Gly Ser Val Asp Phe Lys Lys Asa	
225 230	
TGG ATT CAA TAT AAA GAA GGA TTT GGA CAT CTG TCT CCT ACT GGC ACA	5838
Trp Ile Glu Tyr Lys Glu Gly Phe Gly His Leu Ser Pro Thr Gly Thr	
235 240 245	
ACA GAA TTT TGG CTG GGA AAT GAG AAG ATT CAT TTG ATA AGC ACA CAG	5886
Thr Glu Phe Trp Leu Gly Asa Glu Lys Ile His Leu Ile Ser Thr Glu	
250 255 260 265	
TCT GCC ATC CCA TAT GCA TTA AGA GTG GAA CTG GAA GAC TGG AAT GGC	5934
Ser Ala Ile Pro Tyr Ala Leu Arg Val Glu Leu Glu Asp Trp Asa Gly	
270 275 280	
AGA ACC AG GTACTGTTTT GAAATGACTT CCAACTTTTT ATTGTAAAGA	5982
Arg Thr Ser	
TTGCCTGGAA TGTGCACCTT CCAACTATCA ATAGACAATG GCAAAATGCAG CCTGACAAAT	6042
GCAAAACAGCA CATCCAGCCA CCATTTTCTC CAGGAGTCTG TTTGGTTCTT GGGCAATCCA	6102
AAAAGGTAAA TTCTATTCAG GATGAATCTA AGTGTATTGG TACAATCTAA TTACCCTGGA	6162
ACCATTTCAGA GTAATAGCTA ATTACTGAAC TTTTAATCAG TCCCAGGAAT TGAACATAAA	6222
ATTATAATTT TATCTAGTCT AAATTACTAT TTCATGAAGC AGGTATTATT ATTAATCCCA	6282
TTTTATAGAT TAACTTGCTC AAAATCACAT TGCTGATAAG TGGTAGAGGT AGAATTTCAGA	6342
CTCAAGTAAT TTAACTTTAG AGCCTGTCTT CTIAACAACAT ATCCTGGTTG AAAAGCAAAAT	6402
ACAGCCTCTT CAGACTTCTC AGTGCCCTTG TGGCCATTTA TTCTGTCAAA TCATGAGCTA	6462
CCCTAAAAAT AAACCAAGCTA GCTCTTTTGA TGATCTAGAG GCTTCTTTTT GCTTGAGATA	6522
TTTGAAGGTT TTAAGCATTG TTACCTAATT AAAATGCAAG AAAATATCCA ACCCTCTTGT	6582
TATGTTTTAG GAATAGTGAA ATATATTGTC TTCAAACACA TGGACTTTTT TTTATTGCTT	6642
GGTTGGTTTT TAATCCAGAA AGTGCTATAO TCAGTAGACC TTCTTCTAGG AAAAGGACCTT	6702
CCATTTCCCA GCCACTGGAG ATTAGAAAAAT AAGCTAAATA TTTTCTGGAA ATTTCTGTTT	6762
ATTCATTAAO GCCCATCTT TCCCCCACTC TATAGAAATG TTGTCCACTT GCACAATTTT	6822
TTCAAGGAAA GAATCTCTCT AACTCCTTCA GTCACATGC TTTGGACCAC ACAAGGAAAG	6882
CTTTGATTGT GTAATGCCCT CAAGAAOCTCT CCTTCTTGCC ACTACCACAC TGATTTOAGG	6942
AAGAAAAATC CTTTACACC TAACCCCTTCA GGTGCTATGA GTGGCTAATG GAAGTGTACC	7002
TCCTTCAAAT TTTGTGCAAT AATTAAAGGT CACTCACTGT CAATACTTT CTGTGATCTA	7062
TGATAATGTO TGTOCAACAC ATAACATTTT AATAAAAAATA GAAAAATATGA AATTAGAGTC	7122
ATCTACACAT CTGGATTTGA TCTTAGAATG AAACAAAGCA AAAAGCATCC AAGTGAAGTC	7182
AATTATTAGT TTTCAAGAGT GCTTCAAAAG CTTCTAAGCC CATCCCCGGA AGTGTAAATG	7242
AGCTGTGAGC TGGTTCACAT ATCTATTGCC TCTTGCCAGA TTTGCAAAAA ACTTCACTCA	7302
ATGAGCAAAAT TTCAGCCTTA AGAAACAAAG TCAAAAAATC CAAGGAAAGCA TCCTACGAAA	7362
GAAGGAAACTT CTGAGATCCC TGAGGAGGAT CAACATGTGA TGGTTGTATT TCCTTCTTCT	7422
CAG T ACT GCA GAC TAT GCC ATG TTC AAG GTG GGA CCT GAA OCT GAC	7468
Thr Ala Asp Tyr Ala Met Phe Lys Val Gly Pro Glu Ala Asp	
285 290 295	
AAG TAC CCG CTA ACA TAT OCC TAC TTC OCT GGT GGG GAT OCT GGA GAT	7516
Lys Tyr Arg Leu Thr Tyr Ala Tyr Phe Ala Gly Gly Asp Ala Gly Asp	
300 305 310	
OCC TTT GAT GGC TTT GAT TTT GGC GAT GAT CCT AOT GAC AAG TTT TTC	7564
Ala Phe Asp Gly Phe Asp Phe Gly Asp Asp Pro Ser Asp Lys Phe Phe	
315 320 325 330	

-continued

ACA TCC CAT AAT GGC ATG CAG TTC AOT ACC TGG GAC AAT GAC AAT GAT	7612
Thr Ser His Asn Gly Met Glu Phe Ser Thr Trp Asp Asn Asp Asn Asp	
335 340 345	
AAO TTT GAA GGC AAC TGT GCT GAA CAG GAT GGA TCT GGT TGG TGG ATG	7660
Lys Phe Glu Gly Asn Cys Ala Glu Glu Asp Gly Ser Gly Trp Trp Met	
350 355 360	
AAC AAG TGT CAC GCT GGC CAT CTC AAT GGA GTT TAT TAC CAA G	7703
Asn Lys Cys His Ala Gly His Leu Asn Gly Val Tyr Tyr Glu	
365 370 375	
GTATGTTTTT CTTTCTTAGA TTCCAAAGTTA ATGTATAAGTG TATACTATTT TCATAAAAAA	7763
TAATAAATAAG ATATGAAGAA ATGAAGAATA ATTTATAAAG ATAATAAGGA TTTTATCATG	7823
TTCTTTATTT CAACATAAGTT CTTTGAAACT GGAAGTGGAT AATACCAAAT TCATGCCTAA	7883
AATTAAGCCCT TCTAAAGAAA TCCACCTGCT GCAAAATATC CAGTAGTTTG GCATTATATG	7943
TGAAACTATC ACCATCATAG CTGGCACTGT GGGTTGTGGG ATCTCCTTTA GACATACAAC	8003
ATAAATGATC TGGATGGATT AACATTACTA CATGGATGCT TGTGACACA TTAACCTGGC	8063
TTCCCATGAG CTTTGTGTCA GATACACGCA GTGAACAAGT GTTTGGAGGA ACAGAATAAA	8123
GAAGAAAGCAA GCACTGGTAA GGGCAGGGGT TTGTGAAAAG TTGAGAGAAO AGACCAATCT	8183
GAAGACAGTA GACACTTATT TTAGGATGGG GGTGGATGA GGAAGCTATA GTTTGCTATA	8243
AGCTTGGAAAT GGTTTGGAAC ACTGGTTTCA CTCACCTACC CAGCAGTTAT GTGTGGGGAA	8303
GCCTTACCBA TGCTAAAAGG TCCATGTTAC AATAATGGCA TTATTTGGAA ATCCCAATGG	8363
TATTCATGA ATAAACCAC TATGAAGATA ATCCCACTCA ACAGACTCTC CGTTGGAGAA	8423
GGACAGCAAC ACCACCCTGG GAAAAGCCAAA CAOTCAGACC AGACCTGTTT AGCATCAATA	8483
GGACTTCCCT ACCATATCTG CTGGGTAGAT GAOTGAAAACC AGTGTTCCAA ACCACTCCGG	8543
GCTTGTAOCA AACCATAATC TCCTCATCTA CCAAGATGAG CAACCTTACC TCCTGATGTC	8603
CTAGCCAATC ACCAACTAGG AAACCTTCCA CAGTTTATTT AAAATAACAG TTTGATTTTC	8663
ACAAATATTT TAAATTGGAG AAACATAACT TATCTTTGCA CTCACAAACC ACATAATGAG	8723
AAAGAACTCT AAGGGAAAAAT GCTTGATCTG TGTGACCCGG GGCOCCTATG CAGAGCTGTA	8783
GTTCATGCCA GTTTGTGCT CTGACAAOCC TTTTACAGAA TTACATGAGA TCTGCTTCCC	8843
TAGGACAAAG AGAAAGCCAAA TCAACAGAGG CTGCACTTTA AAATGGAGAC ATAAATAAAC	8903
ATGCCAGAAC CATTTCCTAA AGCTCCTCAA TCAACCAACA AAATTGTGCT TTCAAATAAC	8963
CTGAGTTGAC CTCATCAAGA ATTTTGTGGC TCCTTCTCTT CTAACCTGCC TGAAGAAAGA	9023
TGTTCCACAO CAOCTGAATC CCGGATGGAT AAOCCTAAGG ACAAGAGCCA ATTAAGGAAAC	9083
TTTGGGTTTC TAOCCTTACT AGTAGTGAAT AAATTTAAAG TGTGGATGTG ACTATGATGC	9143
ACAAGACAGA TGTGTTTTAA TAATATGTTT ATTTTATAAA TTGATATTTT AGGAATCTTT	9203
GGAGATATTT TCAATTAGCA GATAATACTA TAAATTTTAT GTAACCTGCA ATGCACCTTC	9263
TAATAAGACAO CTCTTCATAO ACTTGCAOAO GTAAAAAGAT TCCAGAAATA TGAATATGAC	9323
ATCTACGACT TGTTTTAG GT GGC ACT TAC TCA AAA GCA TCT ACT CCT AAT	9373
Gly Gly Thr Tyr Ser Lys Ala Ser Thr Pro Asn	
380 385	
GGT TAT GAT AAT GGC ATT ATT TGG OCC ACT TGG AAA ACC CGG TGG TAT	9421
Gly Tyr Asp Asn Gly Ile Ile Trp Ala Thr Trp Lys Thr Arg Trp Tyr	
390 395 400	
TCC ATG AAG AAA ACC ACT ATG AAG ATA ATC CCA TTC AAC AGA CTC ACA	9469
Ser Met Lys Lys Thr Thr Met Lys Ile Ile Pro Phe Asn Arg Leu Thr	
405 410 415	
ATT GGA GAA GGA CAG CAA CAC CAC CTG GGG GGA GCC AAA CAG GTC AGA	9517

-continued

Ile	Gly	Glu	Gly	Gln	Gln	His	His	Leu	Gly	Gly	Ala	Lys	Gln	Val	Arg	
420					425					430					435	
CCA	GAG	CAC	CCT	GCG	GAA	ACA	GAA	TAT	GAC	TCA	CTT	TAC	CCT	GAG	GAT	9565
Pro	Glu	His	Pro	Ala	Glu	Thr	Glu	Tyr	Asp	Ser	Leu	Tyr	Pro	Glu	Asp	
			440					445					450			
GAT	TTG	TAGAAAATTA	ACTGCTAACT	TCTATTGACC	CACAAAAGTTT	CAGAAAATTCT										9621
Asp	Leu															
CTGAAAAGTTT	CTTCCTTTTT	TCTCTTACTA	TATTTATTGA	TTTCAAOTCT	TCTATTAAGG											9681
ACATTTAGCC	TTCAATGGAA	ATTAAAACTC	ATTTAGGACT	GTATTTCCAA	ATTACTGATA											9741
TCAGAGTTAT	TIAAAAAATTG	TTTATTTGAG	GAGATAACAT	TTCAACTTTG	TTCTAAATA											9801
TATAATAATA	AAATGATTGA	CTTTATTTGC	ATTTTTATGA	CCACTTGTCA	TTTATTTTGT											9861
CTTCGTAAAT	TATTTTCATT	ATATCAAATA	TTTTAGTATG	TACTTAATAA	AATAOGAGAA											9921
CATTTTAGAG	TTTCAAATTC	CCAGGTATTT	TCCTTGTTTA	TTACCCCTAA	ATCATTCCCTA											9981
TTTAATTCTT	CTTTTTAAAT	GGAGAAAAAT	ATGTCTTTTT	AATATGGTTT	TTGTTTTGTT											10041
ATATATTCAC	AGGCTGGAGA	COTTTAAAAAG	ACCGTTTCAA	AAGAGATTTA	CTTTTTTAAA											10101
GGACTTTATC	TGAACAGAGA	GATATAATAT	TTTTCTTATT	GGACAATGGA	CTTGCAAAAGC											10161
TTCACTTCAT	TTTAAGAGCA	AAAGACCCCA	TGTTGAAAAC	TCCATAACAG	TTTTATGCTG											10221
ATGATAATTT	ATCTACATGC	ATTTCAATAA	ACCTTTTGT	TCCTAAGACT	AGATACATGG											10281
TACCTTTATT	GACCATTAAA	AAACCACCAC	TTTTTGCCAA	TTTACCAATT	ACAATTGGGC											10341
AACCATCAGT	AGTAATTOAG	TCCTCATTTT	ATGCTAAATG	TTATGCCTAA	CTCTTTGGGA											10401
GTTACAAAGG	AAATAGCAAT	TATGGCTTTT	GCCCTCTAGG	AGATACAGGA	CAAATACAGG											10461
AAAATACAGC	AACCCAAACT	GACAATACTC	TATACAAGAA	CATAATCACT	AAGCAGGAGT											10521
CACAGCCACA	CAACCAAGAT	GCATAGTATC	CAAAGTGCAG	CTG												10564

(2) INFORMATION FOR SEQ ID NO:6

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 453 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6

Met	Ser	Trp	Ser	Leu	His	Pro	Arg	Asn	Leu	Ile	Leu	Tyr	Phe	Tyr	Ala	
1				5					10					15		
Leu	Leu	Phe	Leu	Ser	Ser	Thr	Cys	Val	Ala	Tyr	Val	Ala	Thr	Arg	Asp	
			20					25						30		
Asn	Cys	Cys	Ile	Leu	Asp	Glu	Arg	Phe	Gly	Ser	Tyr	Cys	Pro	Thr	Thr	
			35				40					45				
Cys	Gly	Ile	Ala	Asp	Phe	Leu	Ser	Thr	Tyr	Gln	Thr	Lys	Val	Asp	Lys	
	50					55					60					
Asp	Leu	Gln	Ser	Leu	Glu	Asp	Ile	Leu	His	Gln	Val	Glu	Asn	Lys	Thr	
	65				70					75				80		
Ser	Glu	Val	Lys	Gln	Leu	Ile	Lys	Ala	Ile	Gln	Leu	Thr	Tyr	Asn	Pro	
			85					90						95		
Asp	Glu	Ser	Ser	Lys	Pro	Asn	Met	Ile	Asp	Ala	Ala	Thr	Leu	Lys	Ser	
		100					105						110			
Arg	Ile	Met	Leu	Glu	Gln	Ile	Met	Lys	Tyr	Gln	Ala	Ser	Ile	Leu	Thr	
	115					120						125				
His	Asp	Ser	Ser	Ile	Arg	Tyr	Leu	Gln	Glu	Ile	Tyr	Asn	Ser	Asn	Asn	
	130					135					140					

-continued

Gln	Lys	Ile	Val	Asn	Leu	Lys	Glu	Lys	Val	Ala	Gln	Leu	Glu	Ala	Gln	
145					150					155					160	
Cys	Gln	Glu	Pro	Cys	Lys	Asp	Thr	Val	Gln	Ile	His	Asp	Ile	Thr	Gly	
				165					170					175		
Lys	Asp	Cys	Gln	Asp	Ile	Ala	Asn	Lys	Gly	Ala	Lys	Gln	Ser	Gly	Leu	
			180					185					190			
Tyr	Phe	Ile	Lys	Pro	Leu	Lys	Ala	Asn	Gln	Gln	Phe	Leu	Val	Tyr	Cys	
		195					200					205				
Glu	Ile	Asp	Gly	Ser	Gly	Asn	Gly	Trp	Thr	Val	Phe	Gln	Lys	Arg	Leu	
	210					215					220					
Asp	Gly	Ser	Val	Asp	Phe	Lys	Lys	Asn	Trp	Ile	Gln	Tyr	Lys	Glu	Gly	
225					230					235				240		
Phe	Gly	His	Leu	Ser	Pro	Thr	Gly	Thr	Thr	Glu	Phe	Trp	Leu	Gly	Asn	
			245						250					255		
Glu	Lys	Ile	His	Leu	Ile	Ser	Thr	Gln	Ser	Ala	Ile	Pro	Tyr	Ala	Leu	
			260					265					270			
Arg	Val	Glu	Leu	Glu	Asp	Trp	Asn	Gly	Arg	Thr	Ser	Thr	Ala	Asp	Tyr	
		275					280					285				
Ala	Met	Phe	Lys	Val	Gly	Pro	Glu	Ala	Asp	Lys	Tyr	Arg	Leu	Thr	Tyr	
	290					295					300					
Ala	Tyr	Phe	Ala	Gly	Gly	Asp	Ala	Gly	Asp	Ala	Phe	Asp	Gly	Phe	Asp	
305				310						315				320		
Phe	Gly	Asp	Asp	Pro	Ser	Asp	Lys	Phe	Phe	Thr	Ser	His	Asn	Gly	Met	
			325						330					335		
Gln	Phe	Ser	Thr	Trp	Asp	Asn	Asp	Asn	Asp	Lys	Phe	Glu	Gly	Asn	Cys	
			340					345					350			
Ala	Glu	Gln	Asp	Gly	Ser	Gly	Trp	Trp	Met	Asn	Lys	Cys	His	Ala	Gly	
		355					360					365				
His	Leu	Asn	Gly	Val	Tyr	Tyr	Gln	Gly	Gly	Thr	Tyr	Ser	Lys	Ala	Ser	
	370					375					380					
Thr	Pro	Asn	Gly	Tyr	Asp	Asn	Gly	Ile	Ile	Trp	Ala	Thr	Trp	Lys	Thr	
385					390					395				400		
Arg	Trp	Tyr	Ser	Met	Lys	Lys	Thr	Thr	Met	Lys	Ile	Ile	Pro	Phe	Asn	
			405						410					415		
Arg	Leu	Thr	Ile	Gly	Glu	Gly	Gln	Gln	His	His	Leu	Gly	Gly	Ala	Lys	
			420					425					430			
Gln	Val	Arg	Pro	Gln	His	Pro	Ala	Gln	Thr	Gln	Tyr	Asp	Ser	Leu	Tyr	
		435					440					445				
Pro	Glu	Asp	Asp	Leu												
			450													

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10807 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: ovine beta- lactoglobulin

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ACGCGTGTCTC	ACCTGCAGGT	CAACGGATCT	CTGTGTCTGT	TTTCATGTTA	GTACCACACT	60
OTTTTGGTGG	CTGTAGCTTT	CACTACAGT	CTGAAGTCAT	AAAAGCCTGGT	ACCTCCAAGT	120
CTGTTCTCTC	TCAAGATTGT	GTCTGTCTGT	TTGGGTCTTT	AGTGTCTCCA	CACAAATTTT	180

-continued

AGAATTGTTT	GTTCTAGTTC	TGTGAAAAAT	GATGCTGGTA	TTTTGATAAG	GATTGCATTG	240
AATCTGTAAA	GCTACAGATA	TAGTCATTGG	GTAGTACAGT	CACTTTAACA	ATATTAACCTC	300
TTACATCTG	TGAGCATGAT	ATATTTTCCC	CCTCTATATC	ATCTTCAATT	CCTCCTATCA	360
GTTTCTTTCA	TTGCAGTTTT	CTGAGTACAG	GTCTTACACC	TCCTTGGTTA	GAGTCATTCC	420
TCAGTATTTT	ATTCCTTTGA	TACAATTGTG	AATGAGGTAA	TTTTCTTAGT	TTCTCTTTCT	480
GATAGCTCAT	TGTTAGTGTA	TATATAGAAA	AGCAACAGAT	TTCTATGTAT	TAATTTTOTA	540
TCCTGCAACA	GATTTCTATG	TATTAATTTT	GTATCCTGCT	ACTTTACGGA	ATTCATTIAT	600
TAGCTTTTTG	GTGACATCTT	GAGGATTTTC	TGAAGAAAAAT	GGCATGGTAT	GGTAGGACAA	660
GGTGTCAATG	CATCTGCAAA	CAGTGGCAGT	TTTCTTTCTT	CCCTTCCAAC	CTGGATTTCCT	720
TTGATTTCTT	TCTGTCTGAG	TACGACTAGG	ATTCCCAATA	CTATACCGAA	TAAAAAGTGGC	780
AAAGAGTGGAC	ATCCTTGTCT	TATTTTTCTG	ACCTTAGAGG	AAATGCTTTC	AGTTTTTCAC	840
CATTAATTAT	AATGTTTACT	GTGGGCTTGT	CATATGTGGC	CTTCATTATA	TGGAGGTCTA	900
TTCCCTCTAT	ACCCACCTTG	TTGAGAATTT	TTATCATAAA	AGTATGTTGA	ATTTTGTCAA	960
AAGTTTTTCC	TGCATCTATT	GAGATGATTT	TTACTCTTCA	ATTCATTAAT	GATTTTTIAT	1020
CTTCATTTTG	TTAATGATTT	CCATTCTTCA	ATTTGTTAAC	GTGGTATATC	ACATTGATTG	1080
ATTTGTGGAT	ACCTTTGTAT	CCCTGGGATA	AACCTCACTT	GATCATGAGC	TTTCAATGTA	1140
TTTTTGAATT	CACTTTGCTA	ATATTCTGTT	GGGTATTTTT	GCATCTCTAT	TCATCAATGA	1200
TATTGOCCTA	AGAAAAGGTTT	TGTCTGGTTT	TAGTATCAGG	GTGATGCTGG	CCTCATAGAG	1260
AGAGTTTAGA	AGCATTTTCT	CCTCTTTGAT	TTTTCGGAAT	AGTTTGAGTA	GGATAAGTAT	1320
TAACTCTTCT	TTAAATGTTT	GGGGACTTCC	CTGGTGAAGC	GGTGGTTGAG	AATCCGCCTC	1380
AGGGATGTGG	GTTTGATCCC	TGGTCAAGGA	ACCATTAAATA	AGATCCCACA	TGCTGCAGGC	1440
AACAAGCCCC	CAAGCTGCAA	CCACTGAGCT	GCAACCGCTG	CAGTGCCAC	AGGCCACGAC	1500
CAGAGAAAAG	CCACATACAG	CAGGGAAAGAC	CCAGCACAAAC	CGGAAAAAGG	AGTTTGGTGG	1560
AATACAGCTG	TGAAGCCGTC	TGGTCCTGGA	CTCCTGCTTG	AGGGAAATTTT	TTAAAAATTA	1620
TTGATTCAAT	TTCAATTACTG	GTAACTGGTC	TGTTCAATAT	TTCTATTTCT	TCCGGGTTCA	1680
GTCTTGGGAG	ATTGTACATG	CCTAAGGAATG	TGTCCGTTTC	TTCTAGGTTG	TCCATTTTAT	1740
TGGACATOCA	TGGGAGCACA	CAGCACCGAC	CAGCGAGACT	CATGCTGGCT	TCCTGGGGCC	1800
AAGCTGGGGC	CCCAAAGCAG	ATGGCATCCT	AGAAGTGTGT	AAAGCCCACT	GACCCTGCCC	1860
AGCCCCACAA	TTTCATTCTG	AGAAAGTATT	CCTTGGTTCT	GCACCTACAG	GCCCAGGATC	1920
TGACCTGCTT	CTGAAGGAOCA	GGGGTTTTTG	CAAGACGGGG	AGATGCTGAG	AGCCGACGGG	1980
GGTCCAAGTC	CCCTCCCAAG	CCCCCTGTCT	TGGGGCAGCC	CTTGGGAAAG	ATTGCCCCAG	2040
TCTCCCTCCT	ACAAGTGTCA	GTCCCAAGCTG	CCCCAAGGCCA	GAGCTGCTTT	ATTTCCGTCT	2100
CTCTCTCTGG	ATGGTATTCT	CTGGAAGCTG	AAAGTTCCCTG	AAATTATGAA	TAGCTTTGCC	2160
CTGAAGGGCA	TGGTTTGTGG	TCACGGTTCA	CAGGAACCTT	GGAGACCCTG	CAGCTCAGAC	2220
GTCCCGAAGT	TGGTGGCACC	CAGATTTCCT	AAGCTCGCTG	GGGAACAAGG	CGCTTGTTC	2280
TCCCTGGCTG	ACCTCCCTCC	TCCCTGCATC	ACCCAATTCT	GAAAAGCAGAG	CGGTGCTGGG	2340
GTCAAGCCCT	CTGCACTCTA	ACGCGGCTGT	CCAAACCACC	CGTGGTGGTG	TTGGGGGGGC	2400
TACCTATGGG	GAAAGGGCTTC	TACTGCAAGT	GGTGGCCCCC	GTCCCTCTG	AGATCAGGAG	2460
TCCCAAGTCCG	GACGTCAAAAC	AAGCCGAGCT	CCCTCCAGAG	GCTCCAGGGA	GGGATCCTTG	2520
CCCCCCCCCT	GCTGCTTCCA	GCTCCTGGTG	CCGCACCCTT	GAGCCTGATC	TTGTAGACGC	2580

-continued

CTCAGTCTAG	TCTCTGCCTC	COTGTTTACA	COCCTTCTCC	CCATGTCCCC	TCCGTGTCCC	2640
COTTTTCTCT	CACAAAGACA	CCGGACATTA	GATTAGCCCC	TGTTCCAGCC	TCACCTGAAC	2700
AGCTCACATC	TGTAAAGACC	TAGATTCCAA	ACAAAGATTCC	AACCTGAAGT	TCCCAGGTGA	2760
TGTGAGTTCT	GGGGCGACAT	CCTTCAACCC	CATCACAGCT	TGCAATTTCAT	CGCAAAAACAT	2820
GGAACTTGGG	GTTTATCOTA	AAACCCAGGT	TCTTCATGAA	ACACTGAGCT	TCGAGGCTTG	2880
TTGCAAGAAAT	TAAAGGTGCT	AATACAGATC	AGGGCAAGGA	CTGAAAGCTGG	CTAAAGCCTCC	2940
TCTTTCCATC	ACAGGAAAAG	GGGGCCTTGG	GGCGGCTTGA	GGTCTGCTCC	COTGAGTGAG	3000
CTCTTTCTCT	CTACAGTCAC	CAACAGTCTC	TCTGGGAAAG	AAACCAAGAG	CCAAGAGAGCA	3060
AGCCGAGGCT	AGTTTAAGAG	ACCCCTGAAC	CTCCACCCAA	GATGCTGACC	AGCCAAGCGGG	3120
CCCCCTGGA	AGACCTTACA	GTTCAAGGGG	GAAAGAGGGG	TGACCCGCCA	GGTCCCTGCT	3180
ATCAGGAGAC	ATCCCCGCTA	TCAAGAGATT	CCCCACCTT	GCTCCCGTTC	CCCTATCCCA	3240
ATACGCCCCAC	CCCACCCCTG	TGATGAGCAG	TTTAGTCACT	TAGAAATGTCA	ACTGAAAGGCT	3300
TTTGATCCCC	CTTTGCCAGA	GGCACAAGGC	ACCCACAAGC	TGCTGGGTAC	CGACGCCCCAT	3360
GTGGATTCAG	CCAGGAGGCC	TGTCCTGCAC	CCTCCCTGCT	CGGGCCCCCT	CTGTGCTCAG	3420
CAACACACCC	AGCACCAAGCA	TTCCCGCTGC	TCCTGAGGTC	TGCAAGCAGC	TCGCTGTAGC	3480
CTGAGCGGTG	TGGAGGGAAAG	TGTCCTGGGA	GATTTAAAAAT	GTAGAGAGCG	GGAGGTGGGA	3540
GGTTGGGCCC	TGTGGGCTTG	CCCATCCCAC	GTGCTGTCAT	TAGCCCCAGT	GCTGCTCAGC	3600
CGTGCCCCCG	CGCGAGGGGT	CAGGTCACTT	TCCCGTCTTG	GGGTTATTAT	GACTCTTGTC	3660
ATTGCCATTG	CCATTTTTTG	TACCTTAACT	GGGCAAGCAG	TGCTTGCAAG	GGCCTCGATA	3720
CCGACCAAGT	CCTCCCTCGG	AGCTCGACCT	GAACCCCATG	TCACCCCTTG	CCCAGCCTGC	3780
AGAGGGTGGG	TGACTGCAAG	GATCCCTTCA	CCCAAAGGCCA	CGGTACATG	GTTTGGAGGA	3840
GCTGGTGCCC	AAGGCAGAGG	CCACCTTCCA	GGACACACCT	GTCCCCAGTG	CTGGCTCTGA	3900
CCTGTCTTGG	TCTAAGAGGC	TGACCCCGGA	AGTGTTCCTG	GCAGTGCGAG	CCAGCCTTGA	3960
CCCAGAGTCC	AGACACCCAC	CTGTGCCCCC	GCTTCTGGGG	TCTACCAAGGA	ACCCTCTAAG	4020
CCCAGAGGGG	ACTTCTGCT	TGGCCTTGGG	TGGAAAGAGG	CCTCCTATTG	TCCTCOTAGA	4080
GGAAAGCCACC	CGGGGGCCTG	AAGATGAGCC	AAATGGGATT	CGGGGAACCG	COTGGCTTGG	4140
GGCCCAAGCC	GGGCTGGCTG	GCCTGTCATG	CTCCTGTATA	AGGCCCCAAG	CCTGCTGTCT	4200
CAGCCCTCCA	CTCCCTGCAG	AGTCAAGAGG	CACGACCCCA	GGGATATCCC	TGCAAGCCATG	4260
AAATGCTCTC	TGCTTGCCCT	GGGCTTGGCC	CTGCTCTGTG	GCCTCCAAGC	CATCATCGTC	4320
ACCAAGACCA	TGAAAGGCTT	GGACATCCAG	AAAGTTTCAAG	GGTTGGCCGG	GTGGGTGAAT	4380
TGCAAGGGCG	GCAGGGGAGC	TGGGCTTCAG	AGAGCCAAAG	GAGGCTGTGA	COTTGGGTTT	4440
CCATCAATCA	GCTAGGGCCA	CCTGACAAAT	CCCCGCTTGG	GCAAGCTTCAA	CCAAGCGTTC	4500
ACTGTCTTGC	ATTCTGGAAG	CTGAAAGCCC	AAAGTCCAGG	TGTTGGCAGG	GCTGGCTTCT	4560
CCTGCGGGCG	CTCTCTGGGG	AGCAGACGGC	CCTCTTCTCC	AGTCTCTGTC	GCAGCCCTGAT	4620
TTCCTCTTCC	TGTGAGGGCA	CCAGGCTTGC	TGGAAACACG	CCTGCTTGGG	CAGCTTCAAC	4680
CGACCTTTGT	CATCTCTTTA	AAAGCCATGT	CTCCAGAGTC	ATGTGTTGAA	GTCTTGGGGG	4740
TTAGTGGGAC	ACAGTTTCAG	CCCTAAAAAG	GTCTCTCTGC	CCCTCAAATT	TTCCCCACCT	4800
CCAAGCATGT	CTCCCCAAGG	TCCAAATGTT	GCTACATGTG	GGGGGGCTCA	TCTGGGTCCC	4860
TCTTTGGGTT	CAGTGTGAGT	CTGGGGAGAG	CATTCCCCAG	GGTGCAGAGT	TGGGGGGAGT	4920
ATCTCAGGGC	TGCCCAGGCC	GGGGTGGGAC	AGAGAGCCCCA	CTGTGGGGCT	GGGGGGCCCC	4980

-continued

TCCCAACCCC	AGAATGCAAC	TCAAAGGTCCC	TCTCCAGGTG	GCAGGGAAGT	GGCACTCCTT	5040
GGCTATGGCG	GCCAGCGACA	TCTCCCTGCT	GGATGCCCAAG	AGTCCCCCCC	TGAGAGTGTAT	5100
CGTGGAGGAG	CTGAAGCCCA	CCCCCGAGGG	CAACCTGGAG	ATCCTGCTGC	AGAAATGGTG	5160
GGCGTCTCTC	CCCAACATGG	AACCCCCACT	CCCCAGGGCT	GTGGACCCCC	CGGGGGGTGG	5220
GGTGCAGGAG	GGACCAGGGC	CCCAGGGCTG	GGGAAAGAGG	CTCAGAGTTT	ACTGGTACCC	5280
GGCGCTCCAC	CCAAAGGCTG	CCACCCAGGG	CTTTTTTTTT	TTTTAAACTT	TTATTAATTT	5340
GATGCTTCAG	AACATCATCA	AACAAATGAA	CATAAAACAT	TCATTTTTGT	TTACTTGGAA	5400
GGGGAGATAA	AATCCTCTGA	AGTGGAAATG	CATAGCAAAAG	ATACATACAA	TGAGGCAAGT	5460
ATTCTGAATT	CCCTGTTAGT	CTGAGGATTA	CAAGTGTATT	TGAGCAACAG	AGAGACATTT	5520
TCATCATTTT	TAGTCTGAAC	ACCTCAGTAT	CTAAAATGAA	CAAGAAAGTCC	TGGAAACGAA	5580
GCAATGTGGG	GATAGGCCCG	TGTGAAAGGT	GCTGGGAGGC	AGCAAGACCTG	GGTCTTCCGG	5640
CTCAAAGCAGT	TCCCCTTACC	AGCCCTGTCC	ACCTCAGAGC	GGGGTCAGGG	TGCAGGAGAG	5700
AGCTGGATGG	GTGTGGGGGG	AGAGATGGGG	ACCTGAACCC	CAGGGCTGCC	TTTTGGGGGT	5760
GCCTGTGGTC	AAGGCTCTCC	CTGACCTTTT	CTCTCTGGCT	TCATCTGACT	TCTCCTGGCC	5820
CATCCACCCG	GTCCCCCTGT	GCCTGAAGGT	ACAGTGAAGT	CGCCGAGGCT	AGTTGGCCAG	5880
CTGGCTCCTA	TGCCCATGCC	ACCCCCCTCC	AGCCCTCCTG	GGCCAAGCTTC	TGCCCCTGGC	5940
CCTCAAGTTCA	TCCTGATGAA	AATGGTCCAT	GCCAATGGCT	CAGAAAGCAG	CTGTCTTTCA	6000
GGGAGAAACG	CGAGTGTGCT	CAGAAAGAGA	TTATTGCAGA	AAAAACCAAAG	ATCCCTGCGG	6060
TGTTCAAAGAT	CGATGGTGAAG	TCCGGGTCCC	TGGGGGACAC	CCACCACCCC	CGCCCCCGGG	6120
GACTGTGGAC	AGGTTCAAGG	GGCTGGCGTC	GGGCCCTGGG	ATGCTAAAGG	ACTGGTGGTG	6180
ATGAAAGACAC	TGCTTGAACA	CCTGCTTCAC	TTGCTTCCCC	TGCCACCTGC	CCGGGGCCTT	6240
GGGGCGGTGG	CCATGGGCAAG	GTCCCCGGCT	GGGGGCTAAC	CCACCAGGGT	GACACCCGAG	6300
CTCTCTTTGC	TGGGGGGCGG	GGGGTGGCTT	GGGCCCTCAG	GCTGAGCTCA	GGAGGTACCT	6360
GTGCCCTCCC	AGGGGTAAAC	GAGAGGCCGT	GGCCACTCCA	GGGGCCCAAG	TGCCCCACGA	6420
CCCCAAGCCCG	CTCCACAAGCT	CCTTCATCTC	CTGAGACAAA	ACTCTGTCCG	CCCTCGCTCA	6480
TTCACTTGTT	CGTCCTAAAT	CCGAGATGAT	AAAGCTTCGA	GGGGGGGTTG	GGGTTCCATC	6540
AGGGCTGCCC	TTCCGGCCGG	CAGCCGTGGG	CACATCTGCC	CTTGGCCCCC	TCAGGACTCA	6600
CTCTGACTGG	AGGCCCTGCA	CTGACTGACG	CCAAGGTGCC	CAGCCCAAGG	TCTCTGGCGC	6660
CATCCAGCTG	CACTGGGTTT	GGGTGCTGGT	CCTGCCCCCA	AGCTGGCCCG	ACACCACAGG	6720
CAAGCCGGGG	TGCCCAGTGG	CCTCGGTCAAG	GGTGAAGCCCC	AGCTGGCCCC	GCTCAGGGCT	6780
TGCCCCGACA	ATGACCCCAT	CCTCAGGACG	CACCCCCCTT	CCCTTGGCTGG	GCAATGTCCA	6840
GGCCCCACCC	AGATCGGGGG	AAAGCCCTATT	TCTTGAACAAC	TCCAATCCCT	GGGGGAGGGG	6900
GGCTCAGACT	GAGTGGTGAAG	TGTTCCCAAG	TCCAAGGAGGT	GGTGGAGGGT	CCTGGCGGAT	6960
CCAAGATTGA	CAATGAAGGG	TTCTGGGGCC	CCATGCGGCT	GGCAATGGCA	GCAAGGAAAG	7020
GGAAAGACCA	TTTCAAGGGT	GGGGGATGCC	AGAGGCGCTC	CCCACCCCGT	CTTGGCCGGG	7080
TGGTGAACCC	GGGGGAGGCC	CGCTGGTGGT	GGAGGGTGGT	GGGGGCTGAC	TAGCAACCCC	7140
TCCCCCCCCG	TTGGAACCTA	CTTTTCTCCC	GTCTTGAACG	CGTCCAAGCT	TGAATGAAGAA	7200
CAAAATCCTT	GTGCTGGACA	CCGACTACAA	AAAGTACCTG	CTCTTCTGCA	TGAAAAACAG	7260
TGCTGAAGCC	GAGCAAAAGC	TGGCCTGCCA	GTGCTGGGGT	GGGTGCCAAC	CCTGGCTGCC	7320
CAGGGAGACC	AGCTGCGTGG	TCCTTGGTGC	AACAAGGGGT	GGGGGGTGGG	AGCTTGAATCC	7380

-continued

CCAGGAGGAG	GAGGGGTGGG	GGGTCCCTGA	GTCCCQCCAG	GAGAGAGTGG	TCGCATACCG	7440
GGAGCCAATC	TGCTGTGGGC	CTGTGGGTGG	CTGGGGACGG	GGGCCAGACA	CACAGGCCGG	7500
GAGACGGGTG	GGCTGCAGAA	CTGTGACTGG	TGTGACCCTC	GCATATGGGC	CGGTGGTCAC	7560
TGAATCTAAC	AGCCTTTGTT	ACCGGGGAGT	TTCAATTATT	TCCCCAAATA	AGAACTCAGG	7620
TACAAAQCCA	TCTTTCAACT	ATCACATCCT	GAAAACAAAT	GGCAGGTGAC	ATTTTCTGTG	7680
CCGTAGCAGT	CCCAGTGGGC	ATTTTCAGGG	CCCCGTGTGC	AGGGGGGGGC	GGGCATCGGC	7740
GAGTGGAGGG	TCCTGGCTGT	GTCAAGCCGG	CCAGGGGGAG	GAAAGGGACC	GGACAGCCAG	7800
AGGTGGGGGG	CAGGCTTTCC	CCCTGTGACC	TGCAGACCCA	CTGCACTGCC	CTGGGAGGAA	7860
GGGAGGGGAA	CTAGGCCCAA	GGGGAGGGGC	AGGTGCTCTG	GAGGGCAAAG	GCAGACCTGC	7920
AGACCACCCT	GGGGAGCAGG	GACTGACCCC	CGTCCCTGCC	CCATAGTCAG	GACCCCGGAG	7980
GTGGACAACG	AGGCCCTGGA	GAAATTGAGC	AAAGCCCTCA	AGGCCCTGCC	CATGCACATC	8040
CGGCTTGCTT	TCAACCCGAC	CCAAGTGGAG	GTGAGACACC	CAAGCCCCGC	CCTTCCCCAG	8100
GGCAGGAGCC	ACCCGGCCCC	GGGACGACCT	CCTCCCATGG	TGACCCCCAG	CTCCCCAGGC	8160
CTCCCAAGAG	GAAAGGGTGG	GGTGCAGCAC	CCCCTGGGGG	CCCCCTCCCC	ACCCCTTCCC	8220
AGGCCTCTCT	TCCCGAGGTG	TCCAGTCCCA	TCCTGACCCC	CCCATGACTC	TCCCTCCCCC	8280
ACAGGGCAGT	GCCACGTCTA	GGTGAAGCCC	TGCCGGTGCC	TCTGGGGTAA	GCTGCCTGCC	8340
CTGCCCCACG	TCCTGGGCAC	ACACATGGGG	TAGGGGGTCT	TGGTGGGGCC	TGGGACCCCCA	8400
CATCAGGCCC	TGGGGTCCCC	CCTGTGAAGAA	TGGCTGGGAG	CTGGGGTCCC	TCCTGGGCGAC	8460
TGCAGAGCTG	GCTGGCCCGG	TGCCACTCTT	GTGGGTGACC	TGTGTCTCTG	CCTCACACAC	8520
TGACCTCCTC	CAGCTCCTTC	CAGCAGAGCT	AAAGCTAAAT	GAGCCAGAAAT	GGTACCTAAAG	8580
GGGAGGGTAA	CGGTCTTTCT	CCCGAGGAGG	GGCTGTCTCT	GAAACCAACG	CCATGGAGAG	8640
GCTGGCAAGG	GTCTGGCAAG	TGCCCCAGGA	ATCACAGGGG	GGCCCCATGT	CCATTTACAG	8700
GCCCCGGGAG	CTTGGACTCC	TCTGGGGACA	GACGACGTCA	CCACCQCCCC	CCCCCATCA	8760
GGGGGACTAG	AAGGGACCAAG	GACTGCAGTC	ACCTTTCTCT	GGACCCAGGC	CCCTCCAGGC	8820
CCCTCCTGGG	GCTCCTGCTC	TGGGCAAGCT	CTCCTTCACC	AATAAAGGCA	TAAACCTGTG	8880
CTCTCCCTTC	TGAGTCTTTG	CTGGAGCAGG	GGCAGGGGGT	GGAGAAAGTGG	TGGGGAGGGAA	8940
GTCTGGCTCA	GAGGATGACA	GGGGGGCTGG	GATCCAGGGC	GTCTGCATCA	CAGTCTTTGT	9000
ACAAGTGGGG	GGCCACACAC	ATCACTGGCG	CTCTTTGAAA	CTTTCAAGAA	CCAGGGAGGG	9060
ACTCGGCAAG	GACATCTGCC	AGTTCACTTG	GAGTGTTCAG	TCAACACCCA	AACTCGACAA	9120
AGGACAGAAA	GTGGAAAATG	GCTGTCTCTT	AGTCTAATAA	ATATTGATAT	GAAACTCAAG	9180
TTGCTCATGG	ATCAATATGC	CTTTATGATC	CAGCCAQCCA	CTACTGTCTG	ATCAACTCAT	9240
GTACCCAAAC	GCACTGATCT	GTCTGGCTAA	TGATGAGAGG	TTCCCAATAG	AGAGCTGGCA	9300
AGAGGTGACA	GTGAGAACTG	TCTGCACACA	CAGCAGAGTC	CACCAATCAT	CCTAAGGAGG	9360
TCAATCCTGG	TGTTCAATGG	AGGACTGATG	TTGAAAGCTG	AACTCCAATG	CTTTGGCCAC	9420
CTGATGTGAA	GAGCTGACTC	ATTTGAAAAA	ACCCTGATGC	TGGGAAAAGAT	TGAGGGCAGG	9480
AGGAGAAAGG	GACGACAGAG	GATGAGATGG	TTGGATGGCA	TCACCAACAC	AATGGACATG	9540
GGTTTGGGTG	GACTCCAGGA	GTTGGTGAAT	GACAGGGAGG	CCTGGCGTGC	TACGGAAAGCG	9600
GTTTATGGGG	TCACAAAGAC	TGAGTGAATG	AACTGAGCTG	AACTGAAATG	AAATGAGGTA	9660
TACAGCAAAAG	TGGGGATTTT	TTAGATAATA	AGAAATATACA	CATAACATAG	TGTATACTCA	9720
TATTTTTTATG	CATACCTGAA	TGCTCAGTCA	CTCAGTCTGA	TCTGACTCTG	TGACCTATGG	9780

-continued

ACCGTAGCCT	TCCAGGTTTC	TTCTGTCCAC	AGAATTCTCC	AAGGCAAGAA	TACTGGAGTG	9840
GGTAGCCATT	TCCTCCTCCA	GGGGATCCTC	CCGACCCAGG	GATTGAACCG	GCATCTCCTG	9900
TATTGGCAGG	TGGATTCTTT	ACCACTGTGC	CACCAAGGAA	GCCCCGTGTTA	CTCTCTATGT	9960
CCCACTTAAT	TACCAAAGCT	GCTCCAAGAA	AAAAGCCCTG	TGCCCTCTGA	GCTTCCCGGC	10020
CTGCAGAGGG	TGGTGGGGGT	AGACTGTGAC	CTGGGAACAC	CCTCCCGCTT	CAGGACTCCC	10080
GGGCCACGTG	ACCCACAGTC	CTGCAGACAG	CCGGGTAGCT	CTGCTCTTCA	AGGCTCATT	10140
TCITTAAGAAA	AAACTGAGGT	CTATTTTGTG	ACTTCGCTGC	CGTAACTTCT	GAACATCCAG	10200
TGCGATGGAC	AGGACCTCCT	CCCCAGGCCT	CAAGGGGCTT	AGGGAGCCAG	CCTTCACCTA	10260
TGAGTCACCA	GACACTCGGG	GGTGGCCCCG	CCTTCAGGGT	GCTCACAATC	TTCCCATCGT	10320
CCTGATCAAA	GAGCAAGACC	AATGACTTCT	TAAGAGCAAG	CAGACACCCA	CAGGACACTG	10380
AGGTTACCA	GAGCTGAGCT	GTCTTTTGA	ACCTAAAGAC	ACACAAGCTCT	CGAAGGTTTT	10440
CTCTTTAATC	TGGATTTAAG	GCCTACTTGC	CCCTCAAGAG	GGAAAGACAAT	CCTGCATGTC	10500
CCCAAGACAG	CCACTCGGTG	GCATCCGAGG	CCACTTAAGT	TTATCTGACC	GCACCCCTGA	10560
ATTAATCGGT	CCAAACTGGA	CAAAAACCTT	GGTGGGAAAT	TTCATCCCA	AGGCCTCAAC	10620
CATCCTGCTT	TGACCACCTT	GCATCTTTTT	TTCTTTTATG	TGTATGCATG	TATATATATA	10680
TATATATTTT	TTTTTTTTTC	ATTTTTTGGC	TGTGCTGGCT	GTTCGTTGCA	GTTCGGTGCG	10740
CAGGCTTCTC	TCTAGTTTCT	CTCTAGTCTT	CTCTTATCAC	AGAGCAATCT	CTAGACGATC	10800
GACGCGT						10807

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AATTCCGATC	GACGCGTCA	CGATATACTC	TAGACGATCG	ACGCGTA	47
------------	-----------	------------	------------	---------	----

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

(B) CLONE: BLGAMP3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGGATCCCCT	GCCGGTGCTT	CTGG	24
------------	------------	------	----

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

(B) CLONE: BLGAMP4

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AACGCGTCAT	CCTCTGTGAG	CCAG	24
------------	------------	------	----

-continued

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 10 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6839

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACTACGTAGT

1 0

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 42 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6632

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGACGCGGAT CCTACGTACC TGCAGCCATG TTTTCCATGA GG

4 2

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6627

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AAGGCTTCGG CAAGCTTCAG G

2 1

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6521

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCCAAAGACT TACTTCCCTC TAGA

2 4

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6520

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCATGAACGT CGCGTGGTGG TTGTGCTACC

3 0

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: ZC6519

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACCACGCGAC GTTCATGCTC TAAAACCGTT

30

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: ZC6518

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCTGCGGGAT CCTACGTACT AGGGGGACAG GGAAGG

36

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: ZC6629

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGACCGGAAT TCTACGTACC TGCAGCCATG AAAAAGGATGG TTTCT

45

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: ZC6630

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGACCGGAAT TCTACGTACC TGCAGCCATG AAACATCTAT TATTO

45

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: ZC6625

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GTGAGATTTT CAGATCTTGT C

21

-continued

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: ZC6626

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AAGAATTACT GTGGCCTACC A

21

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: ZC6624

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GCTGCGGAAT TCTACGTACT ATTGCTGTGG GAA

33

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: ZC6514

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CGACGCGGAT CCTACGTACC TGCAGCCATG AGTTGGTCCT TCAC

45

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: zc6517

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTCTCTGGTA GCAACATACT A

21

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: zc6516

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GGGTTTCTAG CCCTACTAGT AG

22

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: x6515

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGGTTTCTAG CCCTACTAGT AG

22

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AAGCTACGCG TCGATCGTCT AGAGTATATC GTCGACGCGT CGATCGG

47
